

1/69

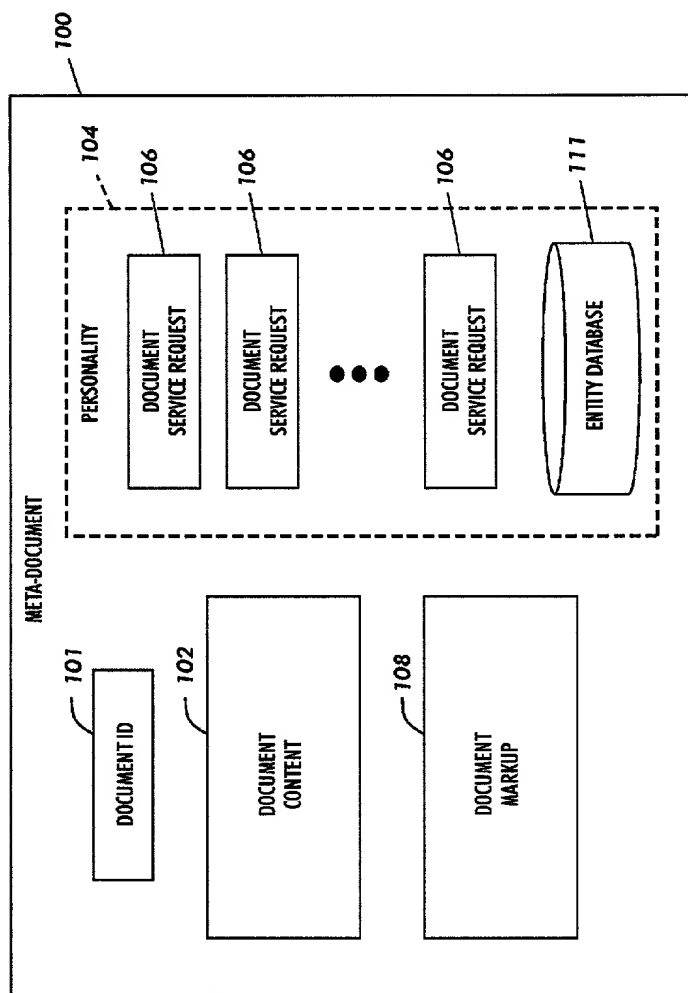


FIG. 1

2/69

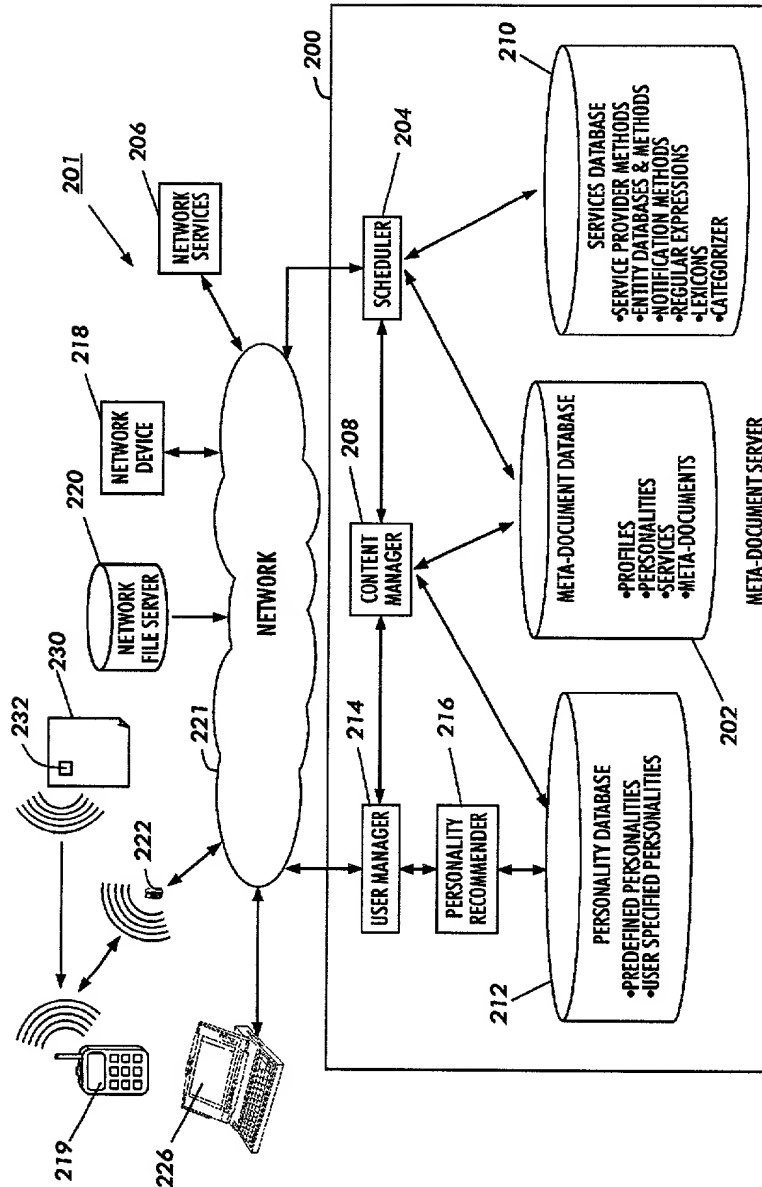


FIG. 2

3/69

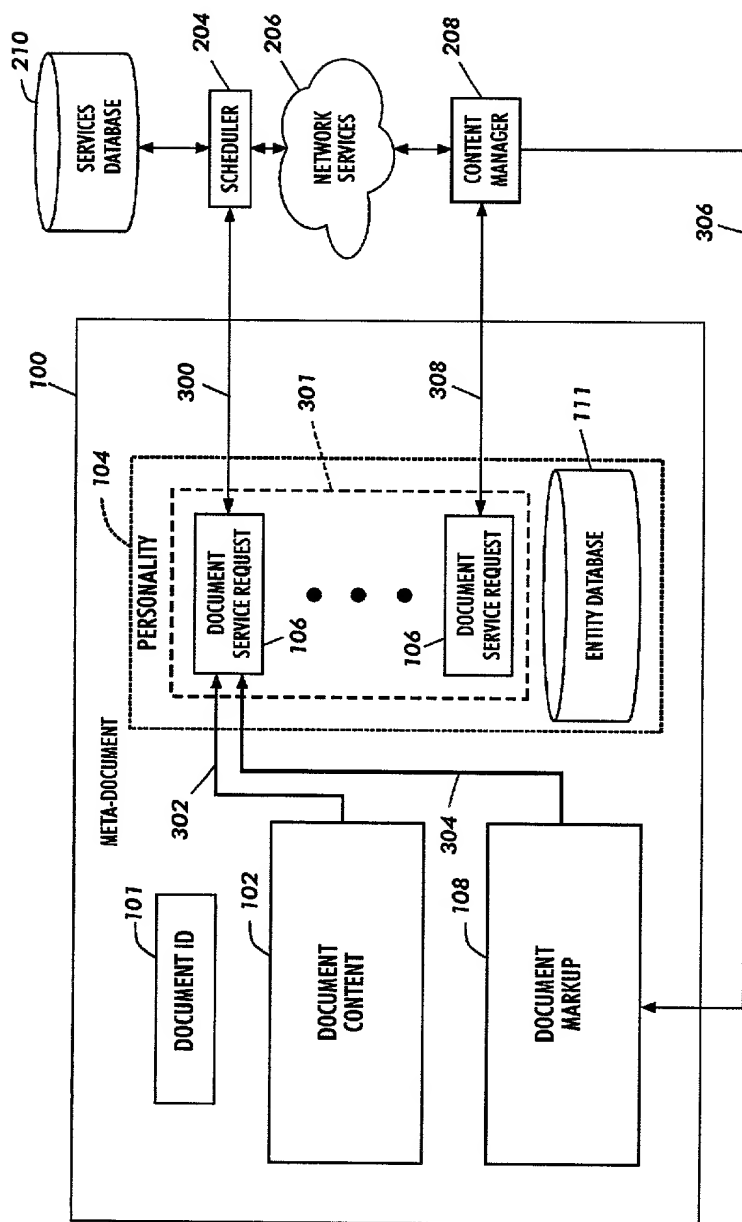
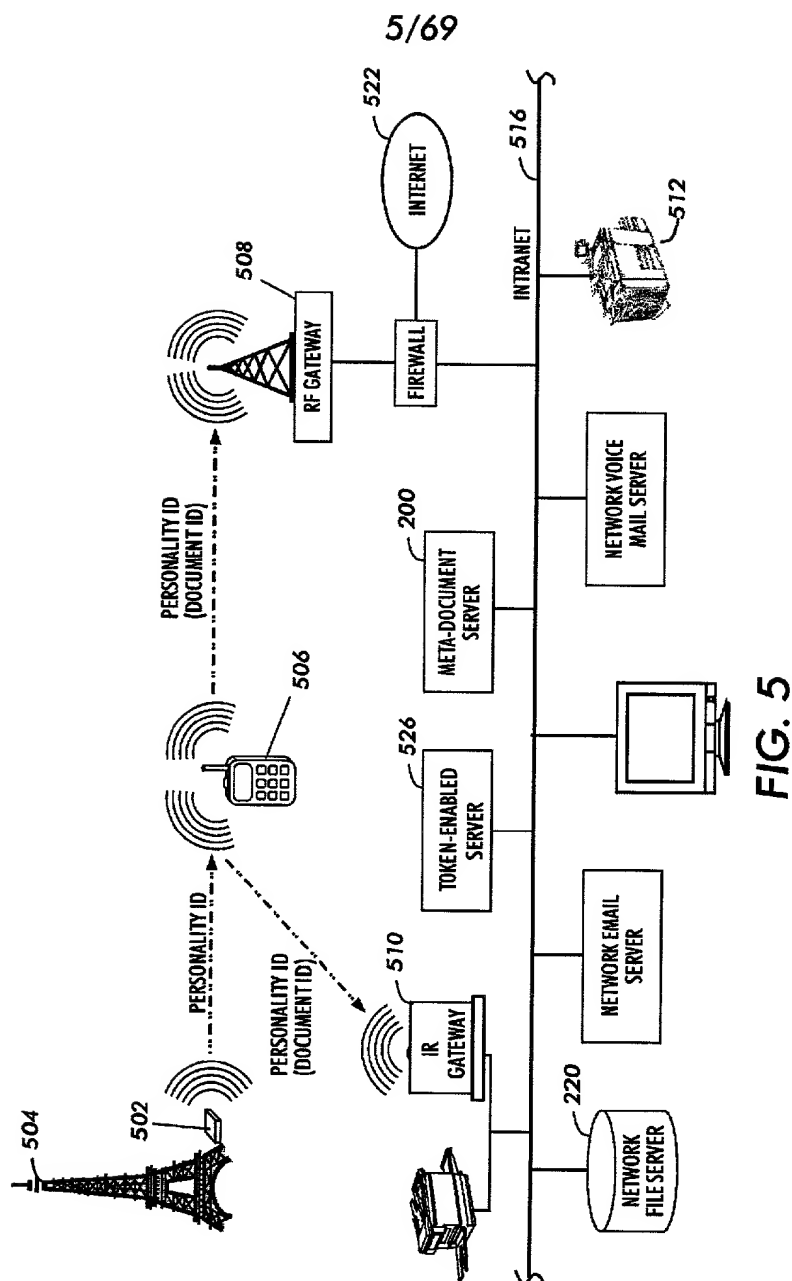
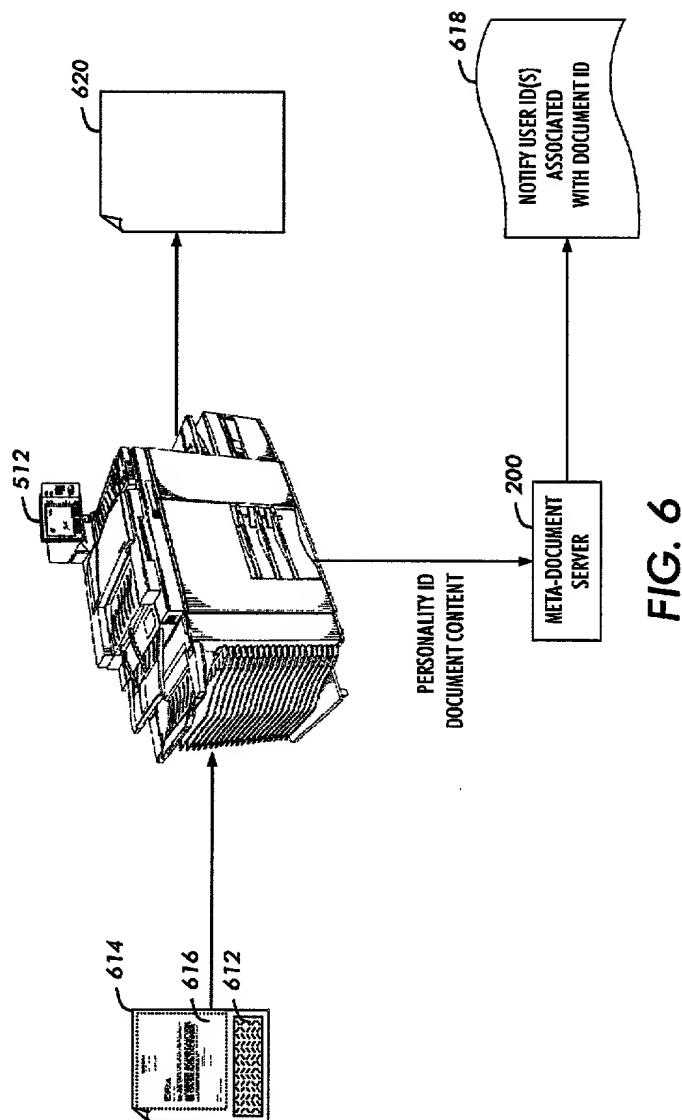


FIG. 3



6/69



7/69

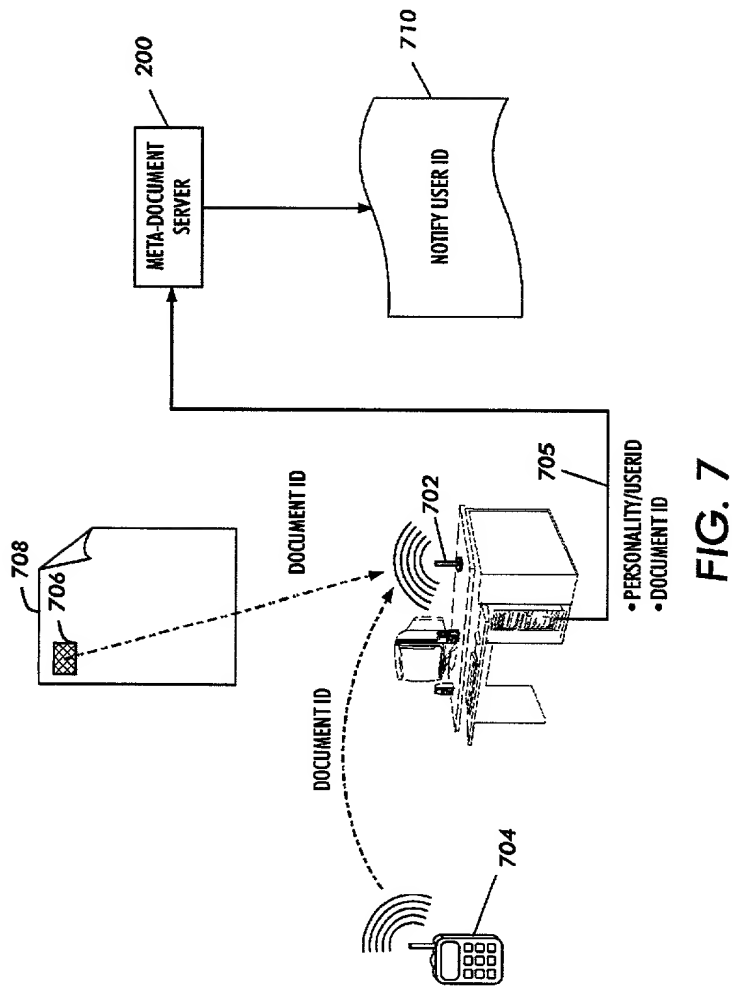


FIG. 7 is a block diagram of a system 700 for document identification. The system 700 includes a document 708 with a marker 706. The marker 706 is associated with a document ID. The document ID is transmitted to a mobile device 704 via a base station 702. The mobile device 704 is associated with a personality/user ID and a document ID. The base station 702 is connected to a meta-document server 200. The meta-document server 200 is connected to a database 710. The database 710 stores the document ID and the personality/user ID.

8/69

PRINT [?] [X]

PRINTER

NAME: 814

STATUS: IDLE

TYPE: XEROX 4517 POSTSCRIPT LEVEL 2

WHERE: TIMES.GRENOBLE.XRCE.XEROX.COM:PASSTHRU ☐ PRINT TO FILE

COMMENT: XEROX N2125 B&W DUPLEX - LE CHATEAU - 2ND FLOOR

PAGE RANGE

☒ ALL

☐ CURRENT PAGE ☐ SELECTION

PAGES:

ENTER PAGE NUMBERS AND/OR PAGE RANGES SEPERATED BY COMMAS. FOR EXAMPLE 1,3,5-12 804

COPIES

NUMBER OF COPIES:

☒ COLLABE 806

ENRICHMENT

☒ PERSONALITY: ☐ SELECT A PERSONALITY FOR ME

APPLY ENRICHMENT TO:

☒ ALL

☐ CURRENT PAGE ☐ SELECTION

PAGES:

INSERT ENRICHMENT AS:

☒ LINKS

☐ CONTENT

808 **ZOOM** **802** **810**

PRINT WHAT: PAGES PER SHEET:

PRINT: SCALE TO PAPER SIZE:

800

FIG. 8

9/69

PRINT DOCUMENT PROPERTIES

LAYOUT PAPER/QUALITY **ENRICHMENT**

902 ☒ PERSONALITY: BUSINESS ☐ SELECT PERSONALITY FOR ME 900

PROVIDE ENRICHMENT BY:

☒ PRINTING WITH DOCUMENT

☐ STORING AT SERVER

INSERT ENRICHMENT AS:

☒ LINKS

☐ CONTENT 901

WHEN STORING ENRICHMENT AT SERVER:

☒ EMAIL ME WHEN COMPLETED

ADDRESS:

☐ DO NOT NOTIFY ME 904

CAPTURE POINT:

☒ AT WORKSTATION 908

☐ AT PRINT SERVER 906

ADVANCED

OK CANCEL

FIG. 9



FIG. 10

11/69

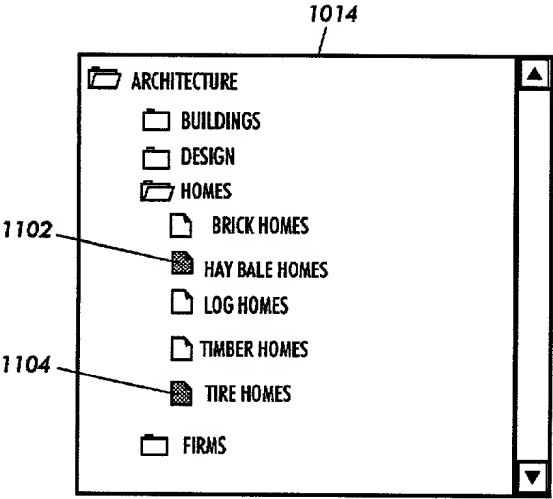


FIG. 11

11/69

12/69

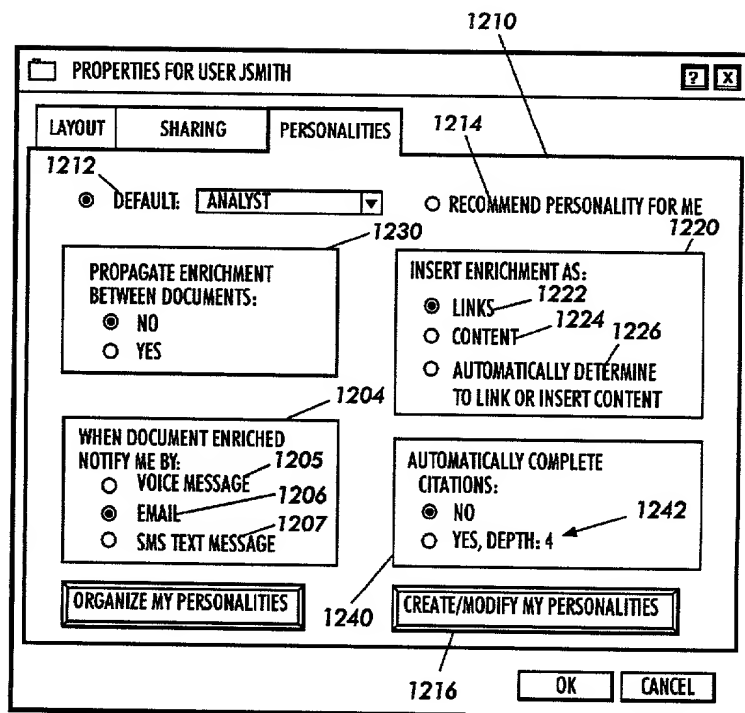


FIG. 12

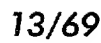


FIG. 13

14/69

1318	1400	1426	1412	1414	1430	1442	1440	1444	1450	
SEARCH ENGINE	NAME	ENTITY TYPE	ANNOTATE	DEPTH	PROVIDER	COST	QUERY	REFRESH		
<input checked="" type="checkbox"/> 1404	COMPUTING	COMPUTING	ALWAYS	0	YAHOO	FREE	LINK	DAILY	1452	
<input type="checkbox"/> 1422	BIOLOGY	BIOLOGY	ALWAYS	0	GOOGLELUCKY	NULL	LINK	NEVER	1451	
<input type="checkbox"/> 1420	MEDICINE	MEDICINE	ALWAYS	0	ALTAVISTA	FEE	LINK	AUTO	1455	
<input type="checkbox"/> 1408	APPLY PAIRS OF ENTITIES AND IDENTIFY 3RD PARTY ENTITIES	COMPANIES	EXPERT	0	GOOGLELUCKY	NULL	CONTENT	NULL	1453	
	APPLY PAIRS OF ENTITIES USING GOOGLELUCKY	PRODUCTS	ALWAYS	2	NORTHERNLIGHTS	NULL	LINK	WEEKLY	1456	
	COUNTRIES	COUNTRIES	NOVICE	0	GOOGLELUCKY	FEE	CONTENT	UNTIL 2003	1454	
	ASTRONOMY	ASTRONOMY	ALWAYS	1	GOOGLE	NULL	LINK	MONTHLY		
	PEOPLE	PEOPLE	ALWAYS	3	GOOGLE:PEOPLE	NULL	LINK	NULL		

1424

VIEW HIERARCHY
SET DEPTH 1432

FIG. 14

15/69

1502

The screenshot shows a window titled "CREATE/MODIFY MY PERSONALITIES" with a standard Windows-style title bar (minimize, maximize, close buttons). The window contains the following elements:

- NAME:** A label followed by a "BROWSE" button and a text input field (labeled 1504).
- Radio Buttons:** Four radio buttons are arranged vertically, each followed by a label:
 - ☒ **CREATE/MODIFY A GROUP OF EXISTING PERSONALITIES** (labeled 1506)
 - ☐ **USE SELECTED DOCUMENTS AND/OR FOLDERS TO CREATE A NEW PERSONALITY** (labeled 1508)
 - Below this label is a "BROWSE" button and a text input field (labeled 1509).
 - ☐ **USE SELECTED FILE/WEBSITE TO CREATE A NEW PERSONALITY** (labeled 1510)
 - Below this label is a "BROWSE" button and a text input field.
 - ☐ **CREATE QUESTION FORM USING PERSONALITY LEXICONS** (labeled 1520)
 - Below this label is a "QUESTION:" label and a text input field.
- Hint:** A label followed by a text input field (labeled 1514).
- KNOWLEDGE LEVEL:** A label followed by two radio buttons:
 - ☒ **NOVICE**
 - ☐ **EXPERT**
- Buttons:** "OK" and "CANCEL" buttons (labeled 1512).

There are also some faint, illegible markings on the left side of the page, possibly from a scanning artifact or a watermark.

FIG. 15

16/69

1600

CREATE/MODIFY A GROUP OF EXISTING PERSONALITIES

NAME: TECHNOLOGY WATCH IN ORGANIC CHEMISTRY

OPERATION:

☒ MERGE 1602

☐ SUBTRACT 1604

PERSONALITY A ▼

WITH/FROM

PERSONALITY B ▼

OK CANCEL

FIG. 16

17/69

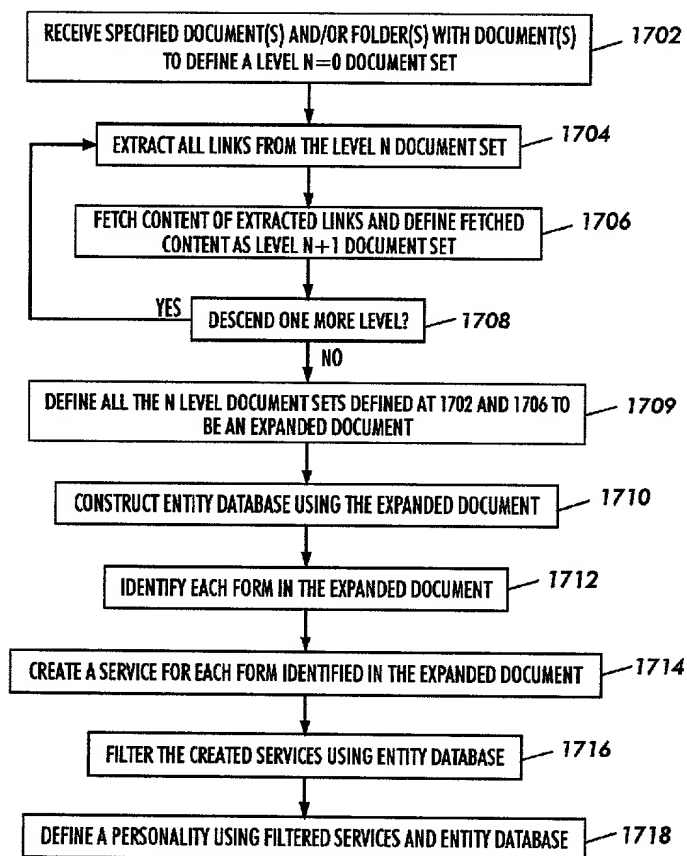


FIG. 17

18/69

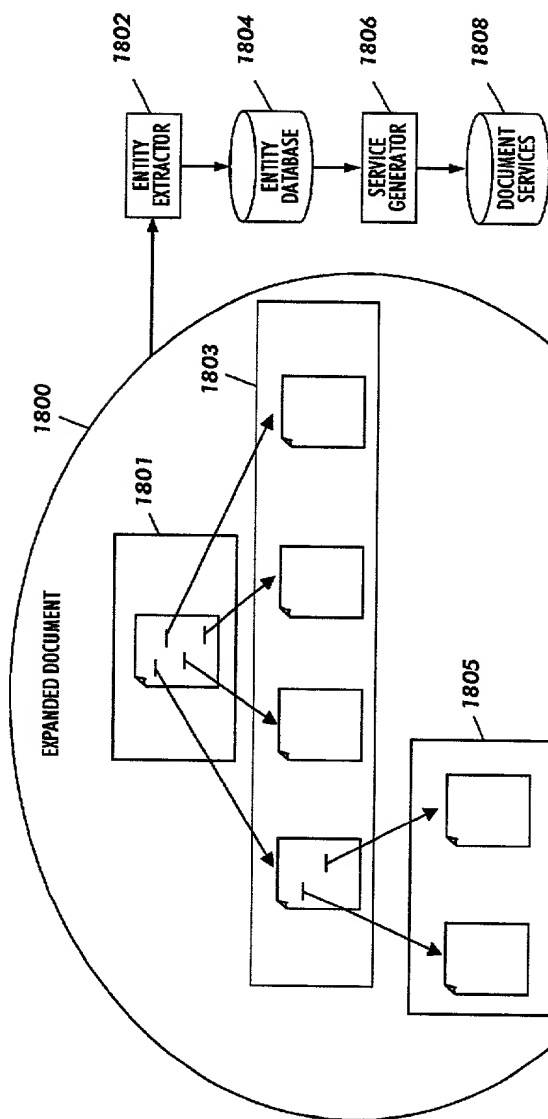


FIG. 18

19/69

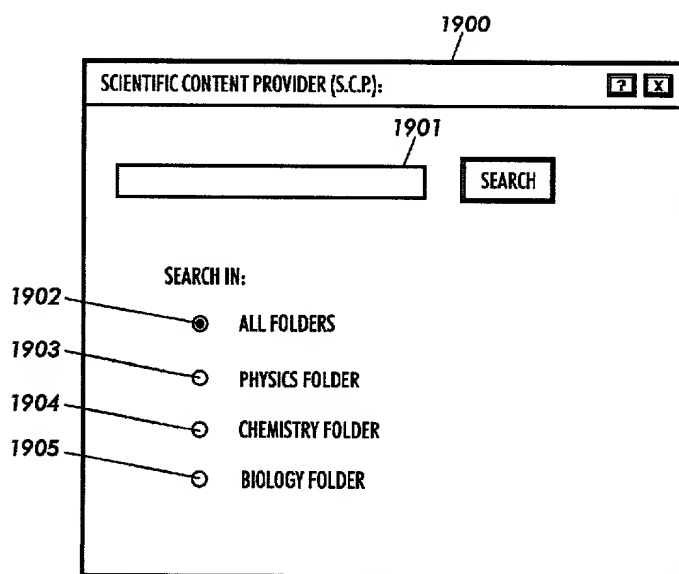


FIG. 19

20/69

SCIENTIFIC SEARCH <input type="checkbox"/> <input type="checkbox"/>									
SEL	NAME		ENTITY TYPE	ANNOTATE	DEPTH	PROVIDER	COST	QUERY	REFRESH
<input checked="" type="checkbox"/>	ALL FOLDERS RESULT OF SCIENTIFIC CONTENT PROVIDER		STANDARD	ALWAYS	0	S.C.P.	FREE	CONTENT	NULL
<input type="checkbox"/>	PHYSICS RESULT OF SCIENTIFIC CONTENT PROVIDER		PHYSICS	ALWAYS	0	S.C.P.	FREE	CONTENT	NULL
<input type="checkbox"/>	CHEMISTRY RESULT OF SCIENTIFIC CONTENT PROVIDER		CHEMISTRY	ALWAYS	0	S.C.P.	FREE	CONTENT	NULL
<input type="checkbox"/>	BIOLOGY RESULT OF SCIENTIFIC CONTENT PROVIDER		BIOLOGY	ALWAYS	0	S.C.P.	FREE	CONTENT	NULL

2004

2002

FIG. 20

21/69

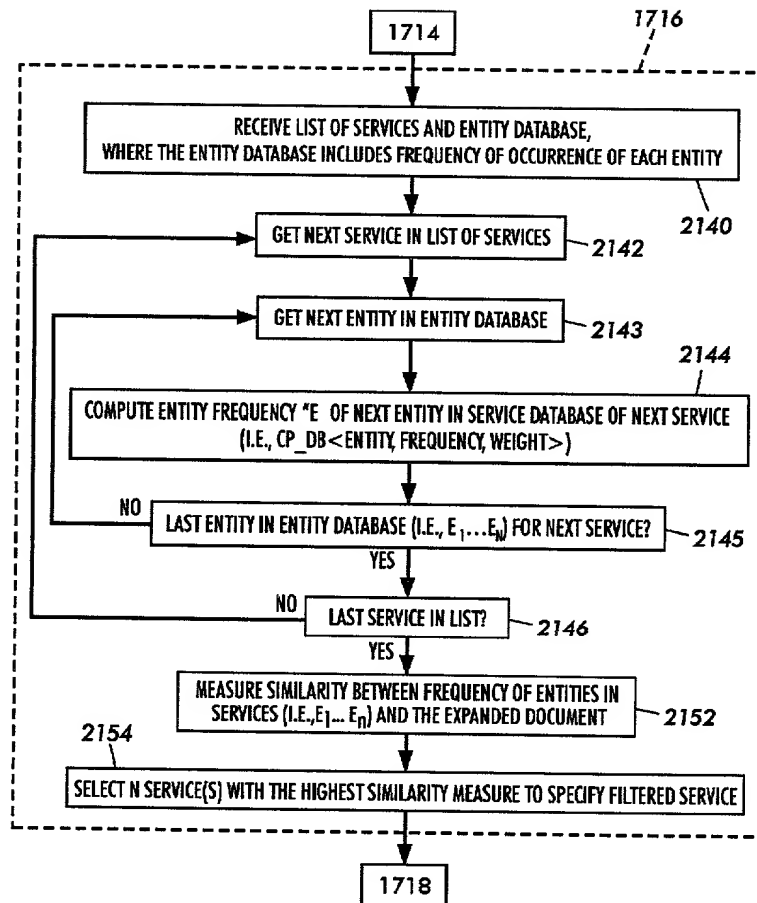
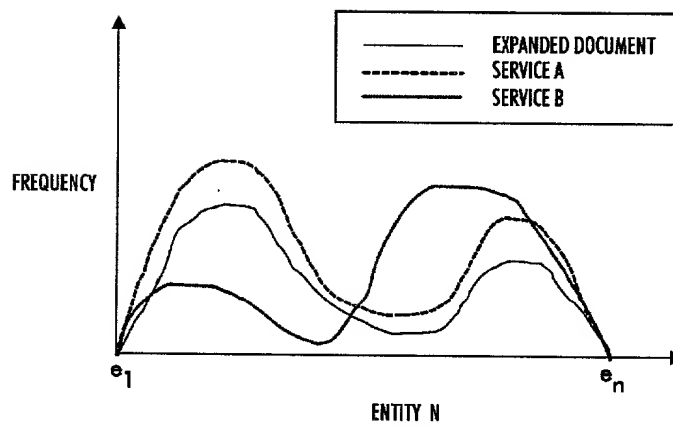


FIG. 21

22/69

**FIG. 22**

23/69

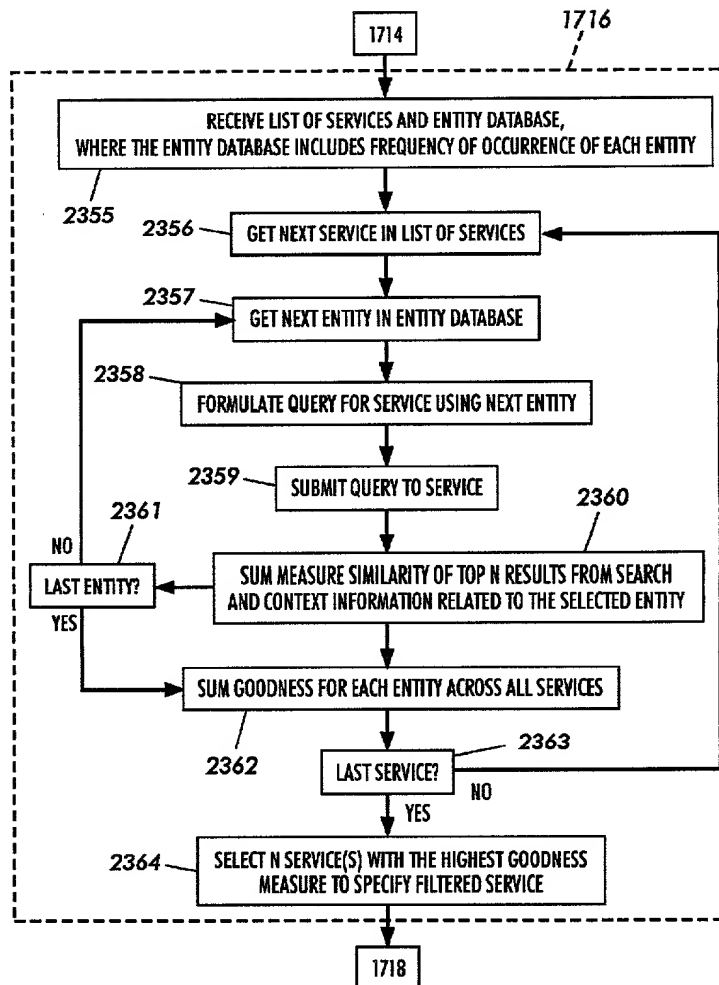


FIG. 23

24/69

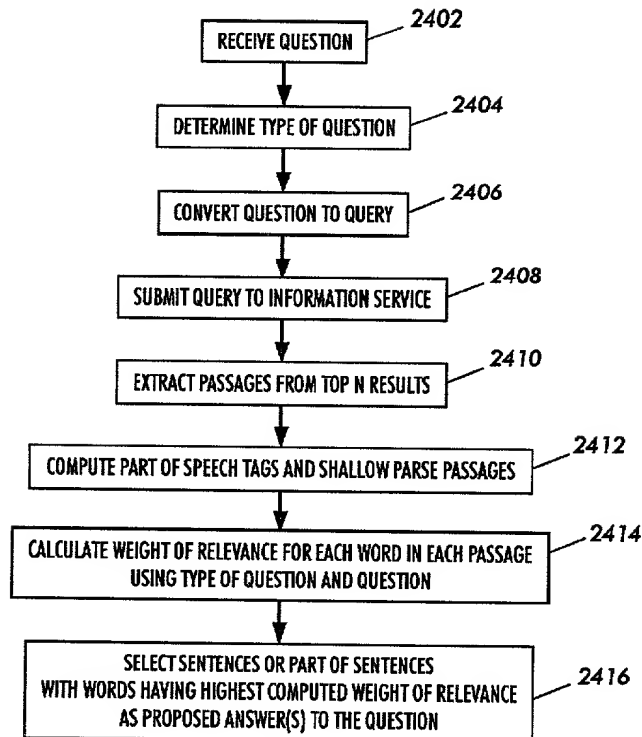


FIG. 24

25/69

2502

E-LEARNING SERVICES ASSOCIATED WITH E-LEARNING PERSONALITY	
SEL	
<input checked="" type="checkbox"/>	LINK WORDS OR MULTWORD EXPRESSIONS FOUND IN THE DOCUMENT TO THEIR DEFINITIONS
<input type="checkbox"/>	LINK WORDS OR MULTWORD EXPRESSIONS FOUND IN THE DOCUMENT TO THEIR TRANSLATIONS
<input type="checkbox"/>	LINK EACH TEXT UNIT (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) TO A TUTORIAL CONCERNING EACH TEXT UNIT
<input type="checkbox"/>	LINK EACH TEXT UNIT (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) TO AN EXPERT CONCERNING IDENTIFIED SUBJECTS
<input type="checkbox"/>	LINK EACH TEXT UNIT (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) TO AN INTERACTIVE COURSE ABOUT IDENTIFIED SUBJECTS
<input checked="" type="checkbox"/>	LINK EACH TEXT UNIT (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) TO A LIST OF REAL WORLD CLASSES ABOUT IDENTIFIED SUBJECTS
<input checked="" type="checkbox"/>	LINK TEXT UNITS TO (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) LEARNING RESOURCES CONCERNING EACH TEXT UNITS

FIG. 25

26/69

2602

LANGUAGE LEARNING SERVICES ASSOCIATED WITH LANGUAGE LEARNING PERSONALITY	
SEL	
<input checked="" type="checkbox"/>	LINK WORDS OR MULTWORD EXPRESSIONS FOUND IN THE DOCUMENT TO THEIR DEFINITIONS
<input type="checkbox"/>	LINK WORDS OR MULTWORD EXPRESSIONS FOUND IN THE DOCUMENT TO THEIR TRANSLATIONS
<input type="checkbox"/>	LINK EACH SENTENCE, OR PHRASE, TO A GRAMMATICAL DESCRIPTION OF THE STRUCTURE OF THE SENTENCE OR PHRASE
<input type="checkbox"/>	LINK EACH WORD, MULTWORD EXPRESSION, PHRASE OR SENTENCE TO OTHER INSTANCES OF THE SAME IN DIFFERENT CONTEXTS FROM THE PRESENT
<input type="checkbox"/>	LINK EACH WORD, MULTWORD EXPRESSION, PHRASE OR SENTENCE TO ONE OR MORE INTERACTIVE GRAMMAR EXERCISES CONCERNING THAT ELEMENT
<input checked="" type="checkbox"/>	LINK EACH TEXT UNIT (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) TO A LIST OF REAL WORLD CLASSES ABOUT IDENTIFIED LANGUAGE
<input checked="" type="checkbox"/>	LINK TEXT UNITS TO (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) LANGUAGE LEARNING RESOURCES CONCERNING EACH TEXT UNITS

2604
2606
2608
2610
2612
2614
2616

FIG. 26

27/69

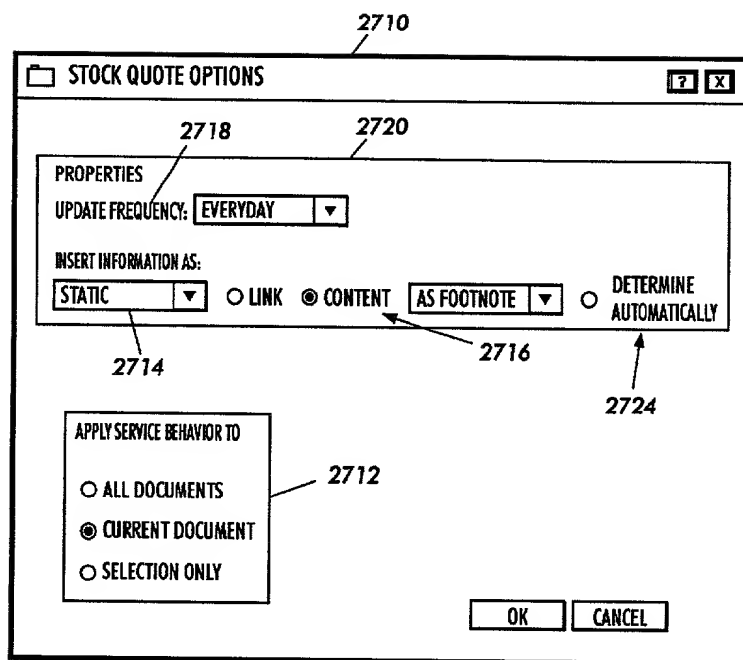


FIG. 27

28/69

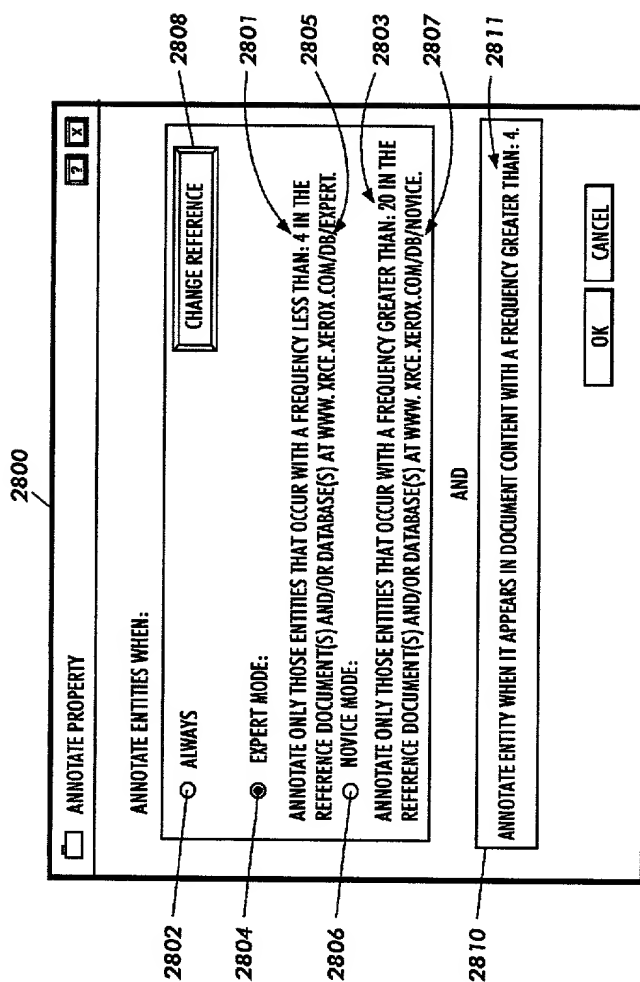


FIG. 28

29/69

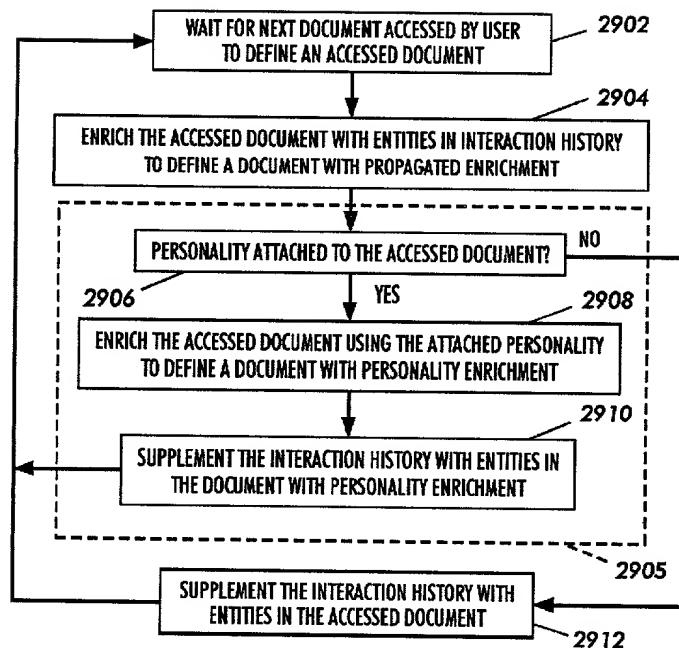


FIG. 29

30/69

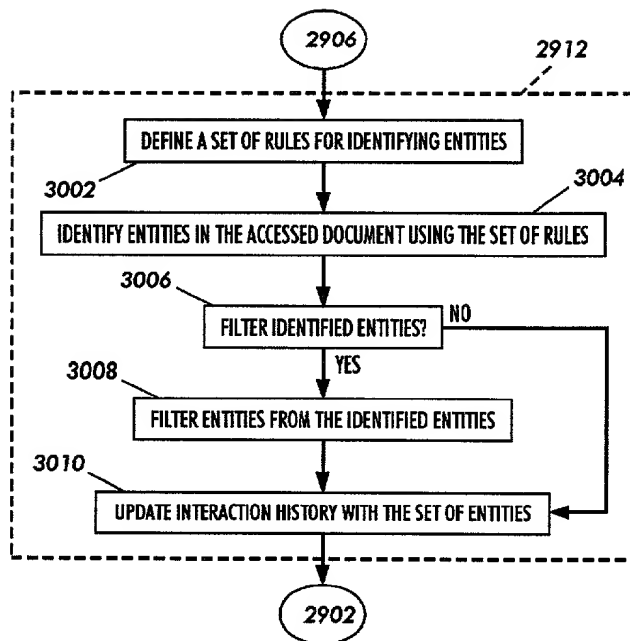


FIG. 30

31/69

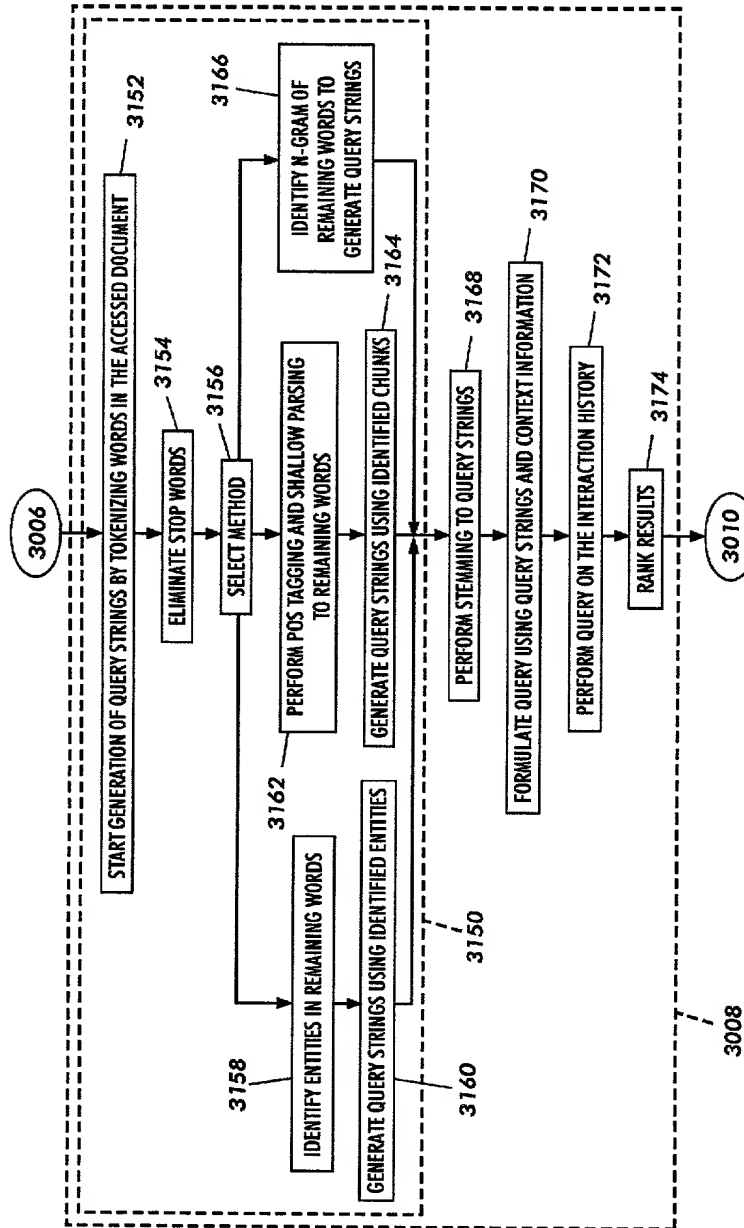


FIG. 31

32/69

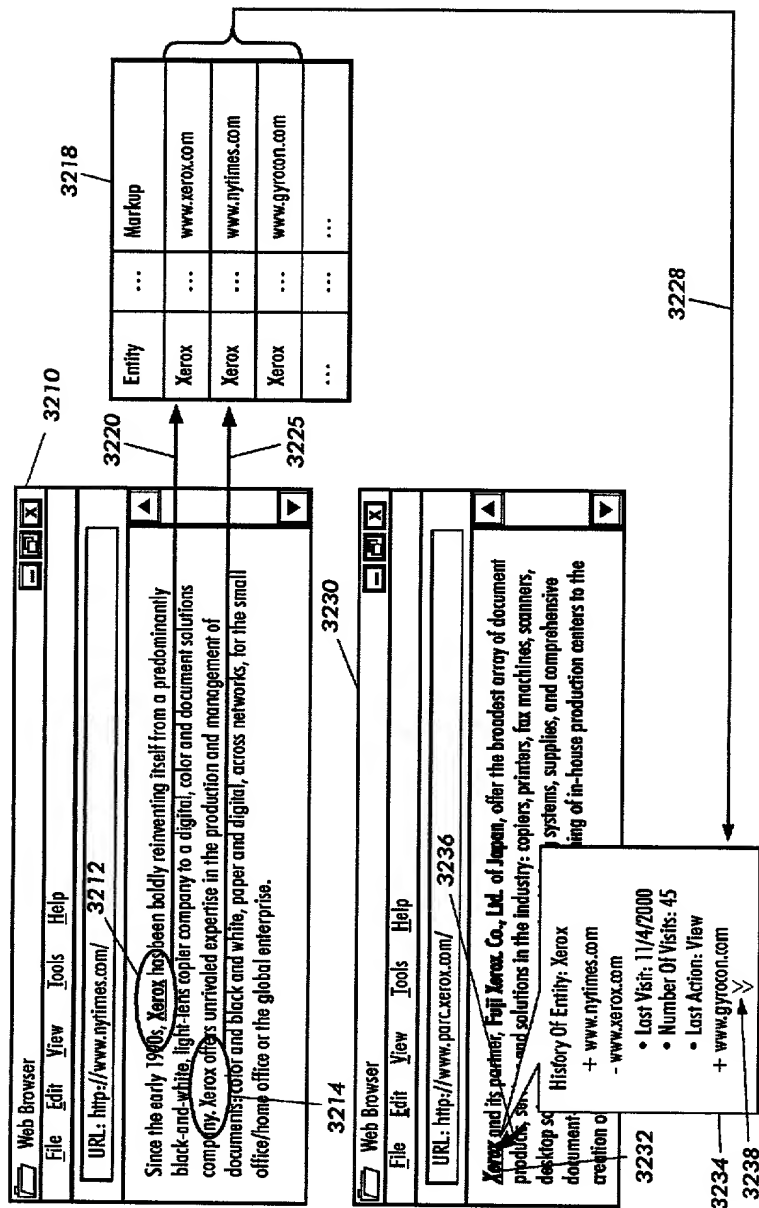


FIG. 32

33/69

3218

ENTITY	ENTITY TYPE	VISITS	LAST VISIT	ACTION LIST	DMOZ CATEGORY	LOCATION	POS	MARKUP
XEROX	BUSINESS	45	11/4/2000	VIEW (43), PRINT(2)	BUSINESS	40	NOUN	WWW.XEROX.COM
XEROX	BUSINESS	1	5/1/2001	VIEW(1)	FINANCE	689	NOUN	WWW.NYTIMES.COM
XEROX	BUSINESS	4	2/22/2001	EMAIL(1), VIEW(3)	COMPUTER	343	NOUN	WWW.GYROCON.COM
...

FIG. 33

34/69

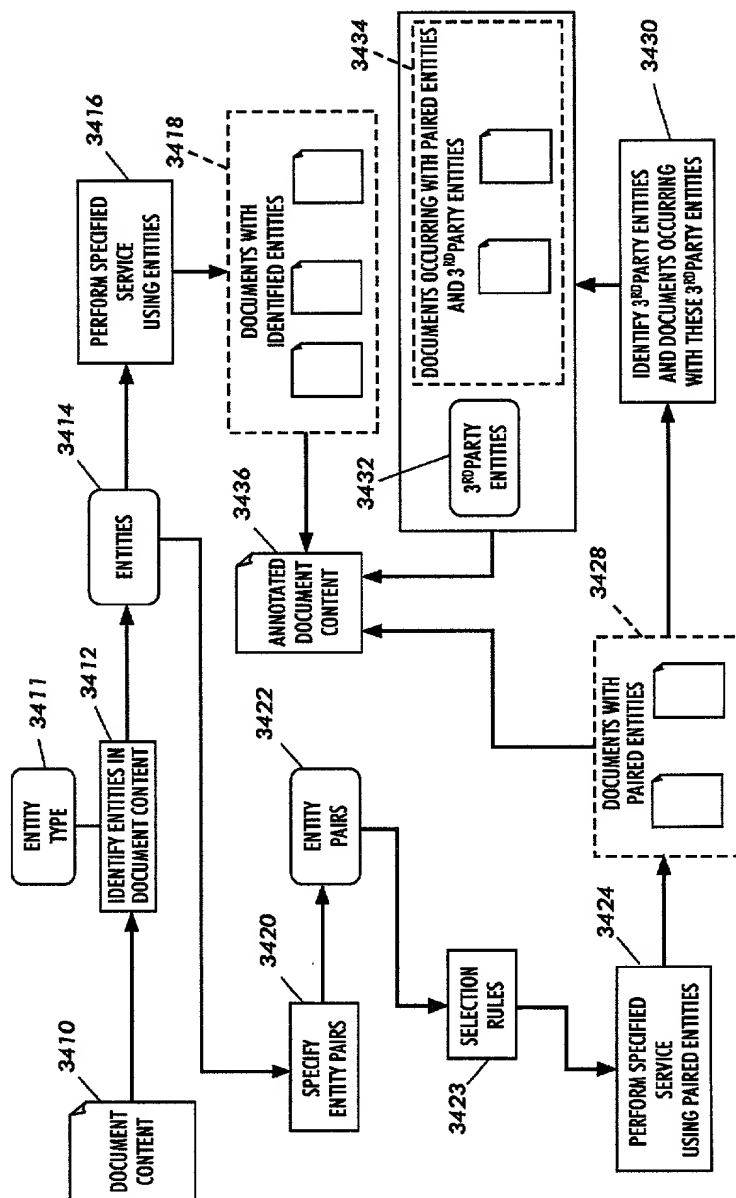


FIG. 34

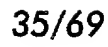


FIG. 35

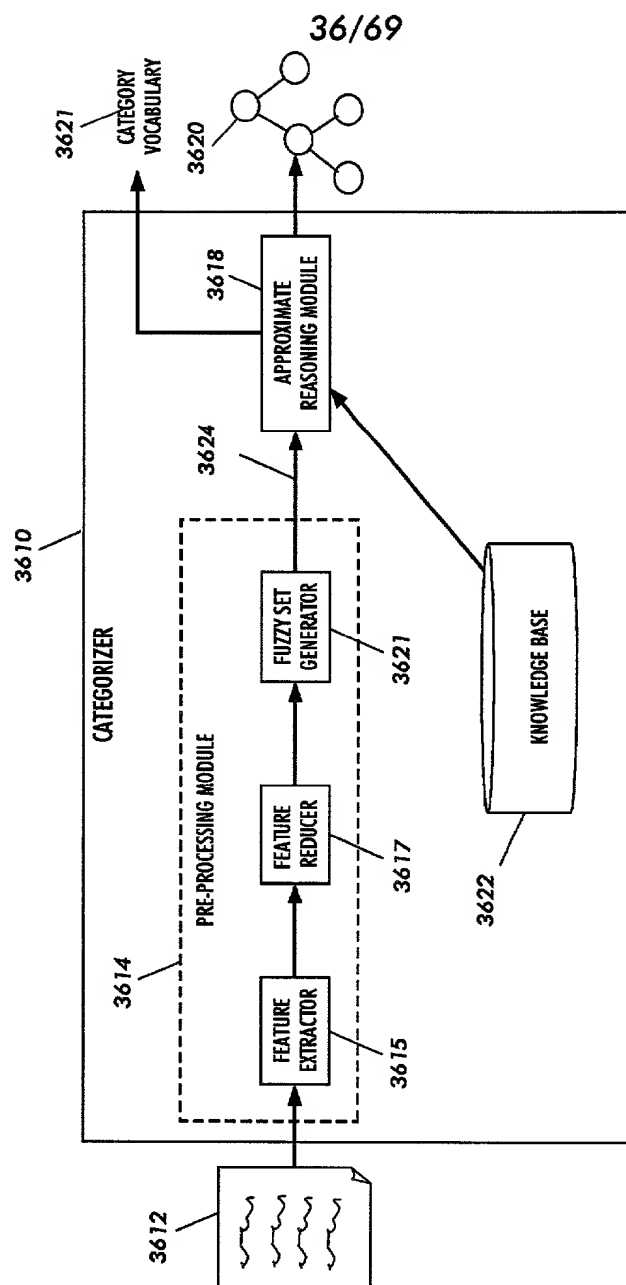
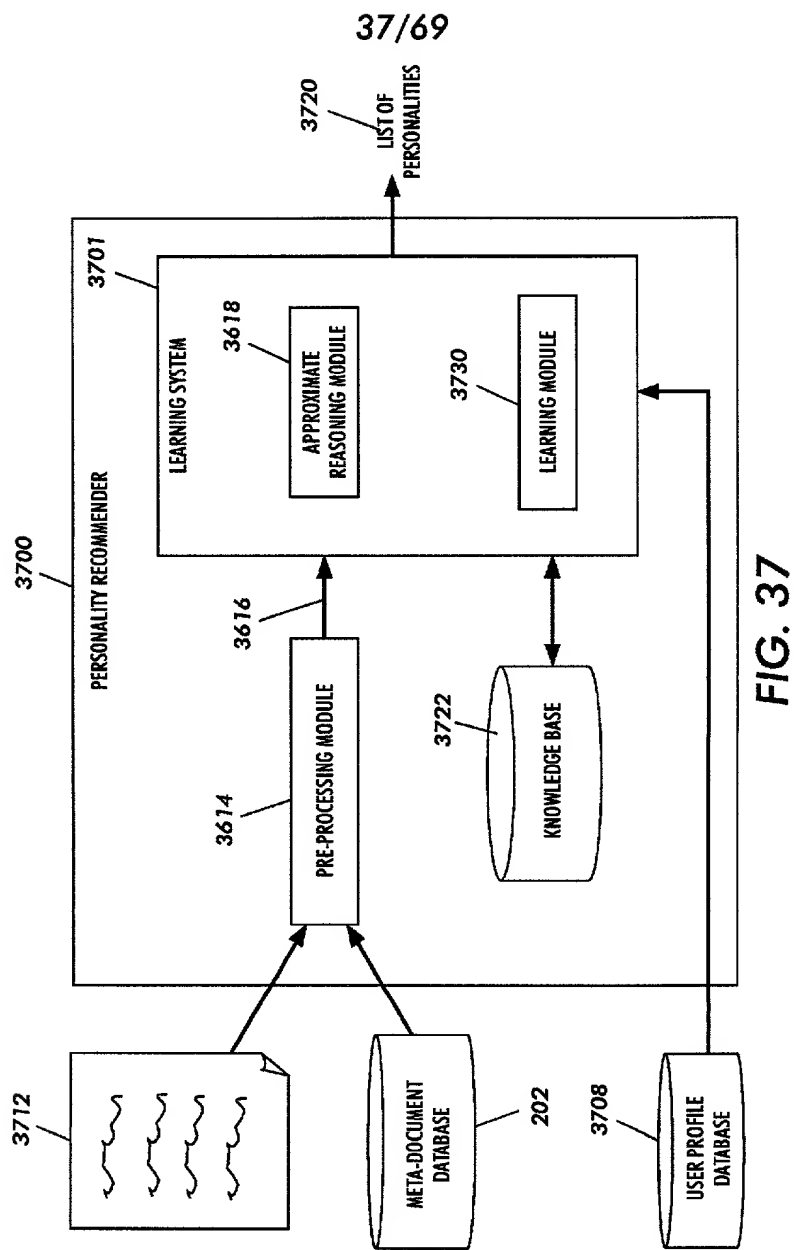


FIG. 36



38/69

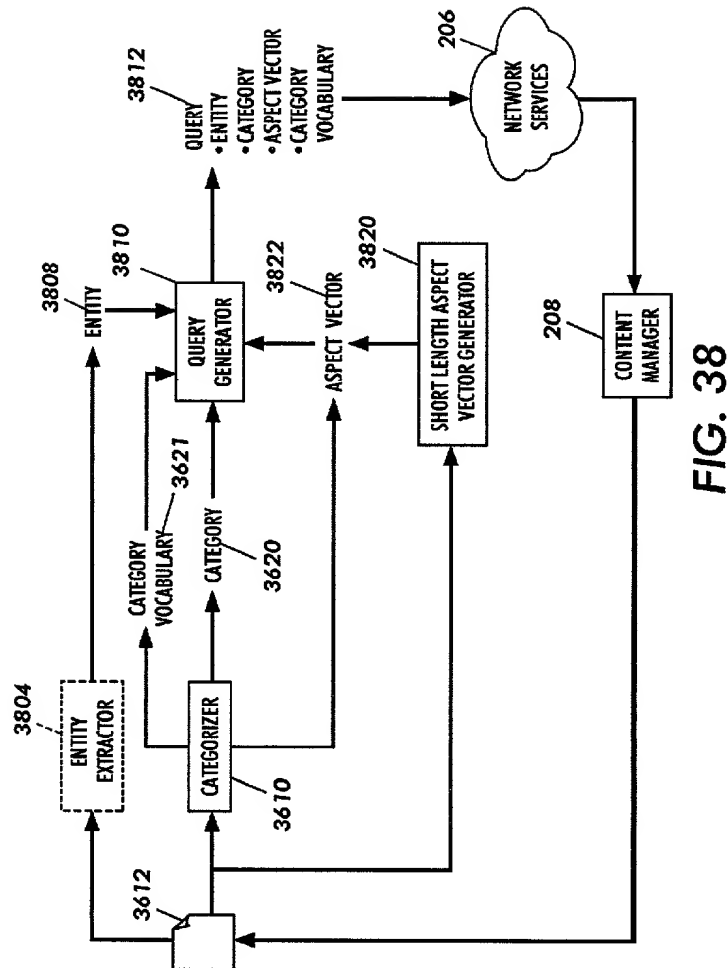


FIG. 38

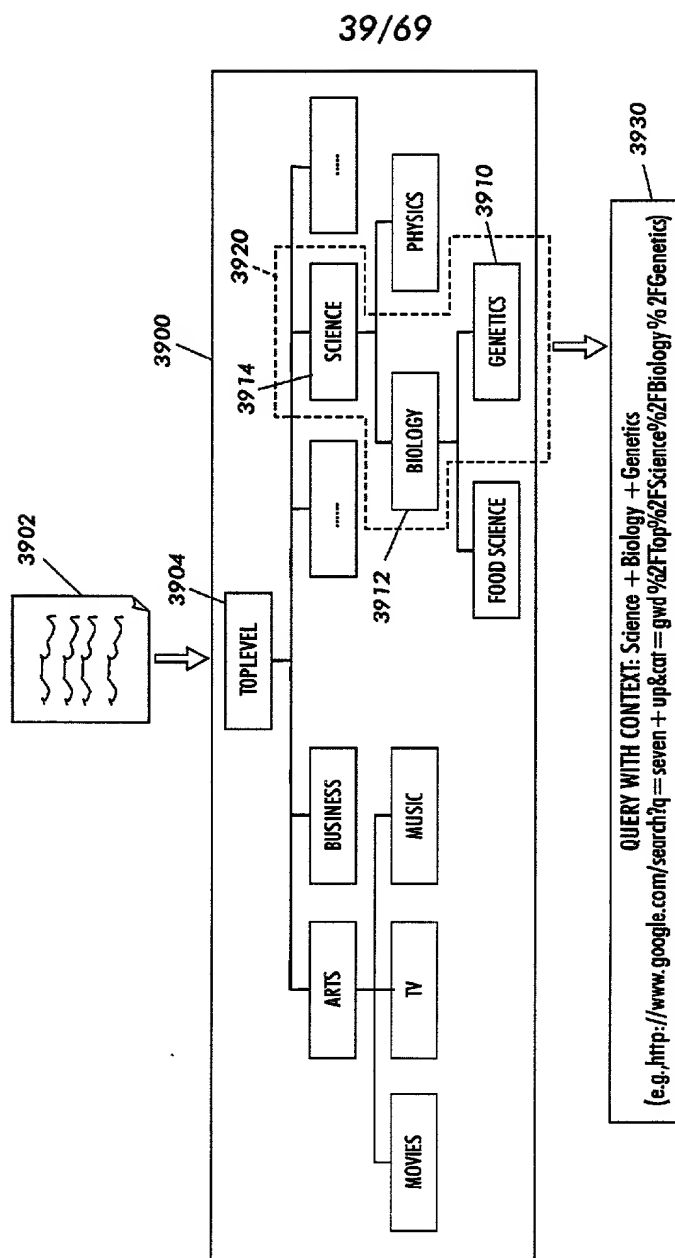


FIG. 39

40/69

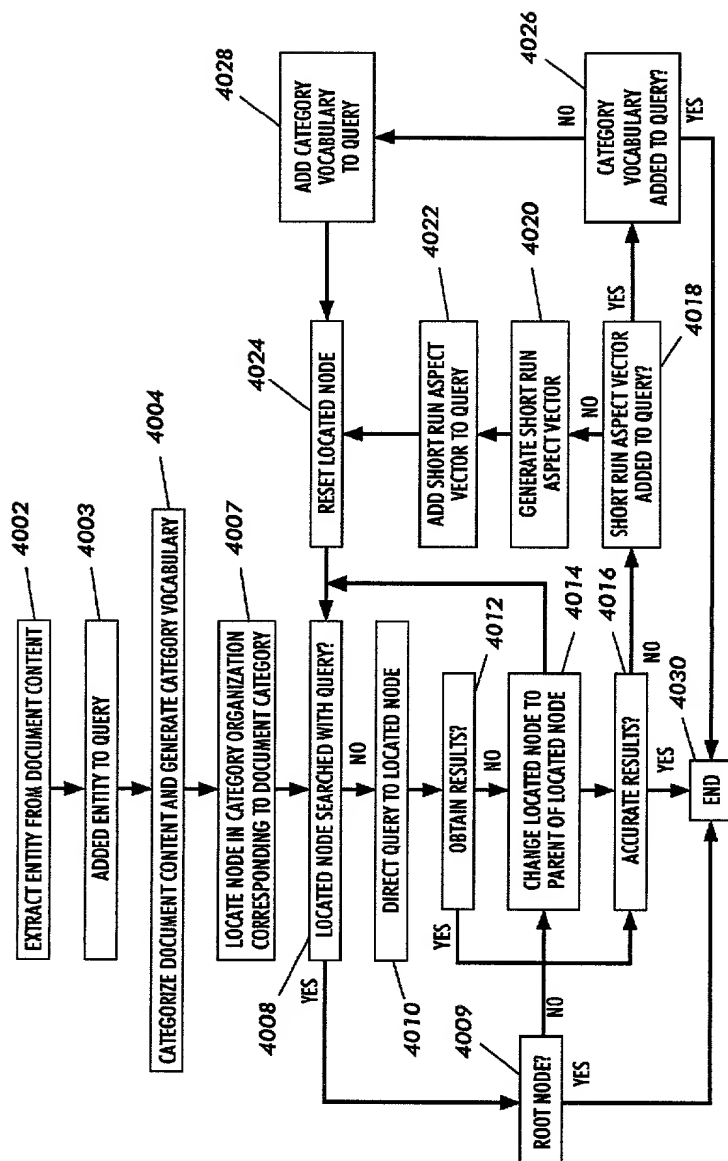


FIG. 40

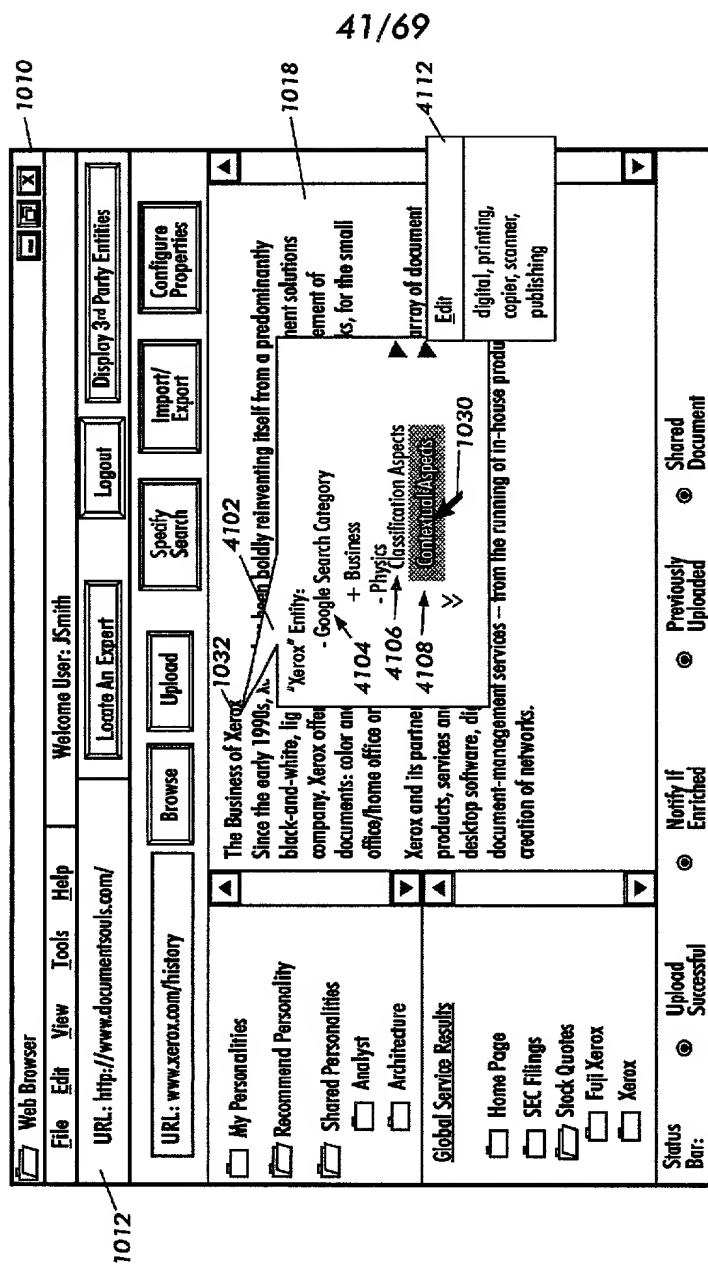


FIG. 41

42/69

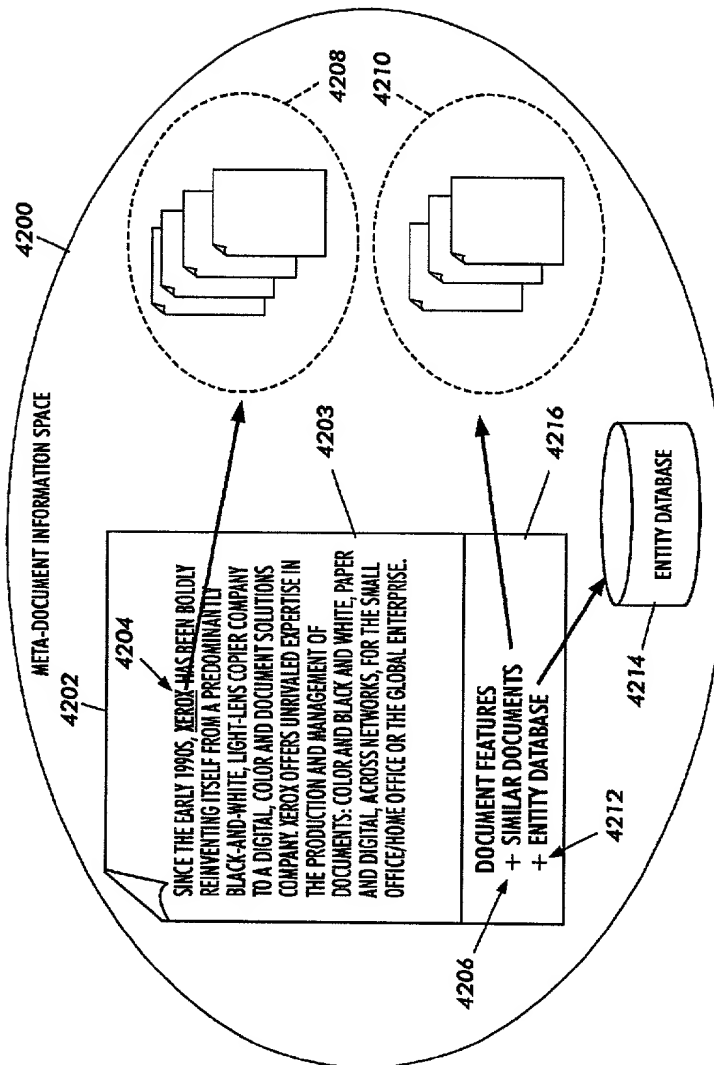


FIG. 42

43/69

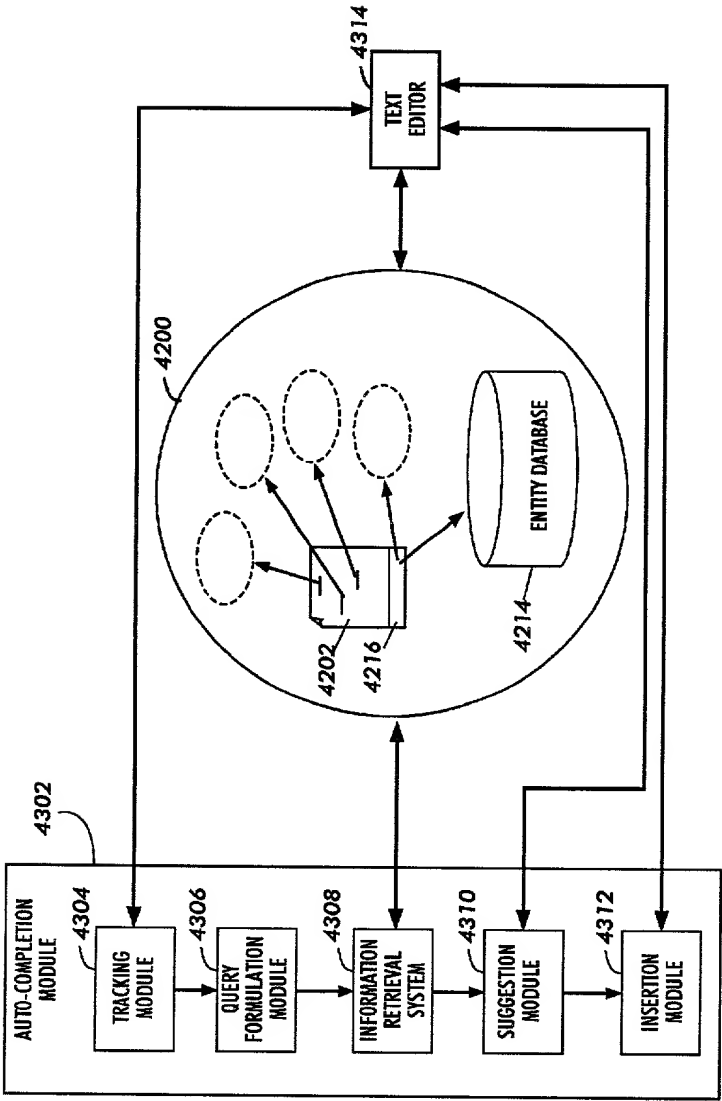


FIG. 43

FIG. 43 is a block diagram of an auto-completion system. The system includes a text editor (4314) and an auto-completion module (4302). The auto-completion module (4302) includes a tracking module (4304), a query formulation module (4306), an information retrieval system (4308), a suggestion module (4310), and an insertion module (4312). The text editor (4314) is connected to the tracking module (4304) and the insertion module (4312). The tracking module (4304) is connected to the query formulation module (4306), which is connected to the information retrieval system (4308), which is connected to the suggestion module (4310), which is connected to the insertion module (4312). The insertion module (4312) is connected back to the text editor (4314). The auto-completion module (4302) is also connected to an entity database (4214) via a bidirectional connection. The entity database (4214) is part of a larger system (4200) that also includes a central processing unit (4202) and several other components (4216).

44/67

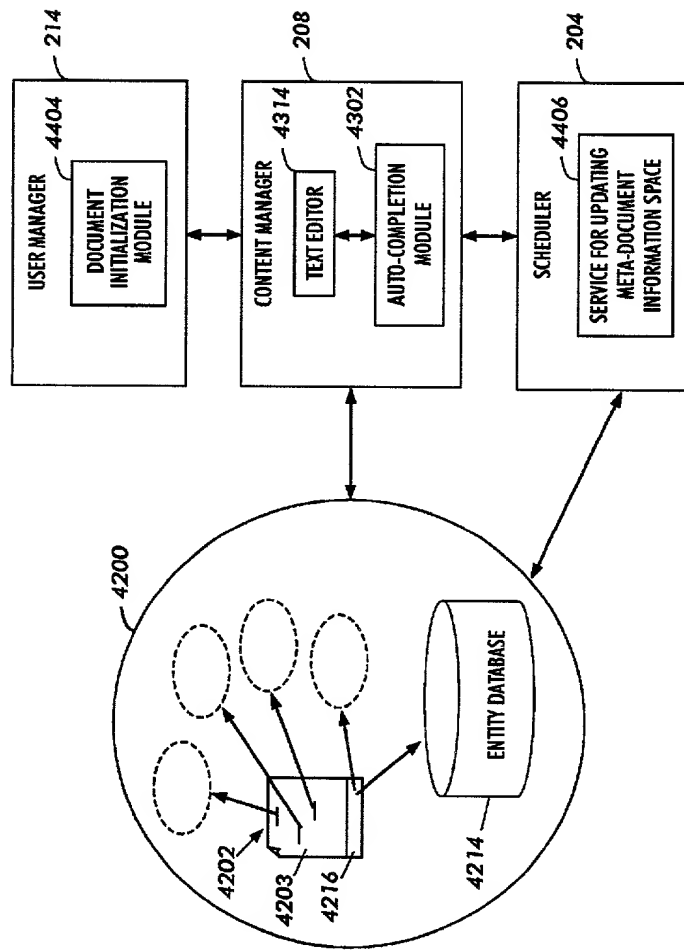


FIG. 44

45/67

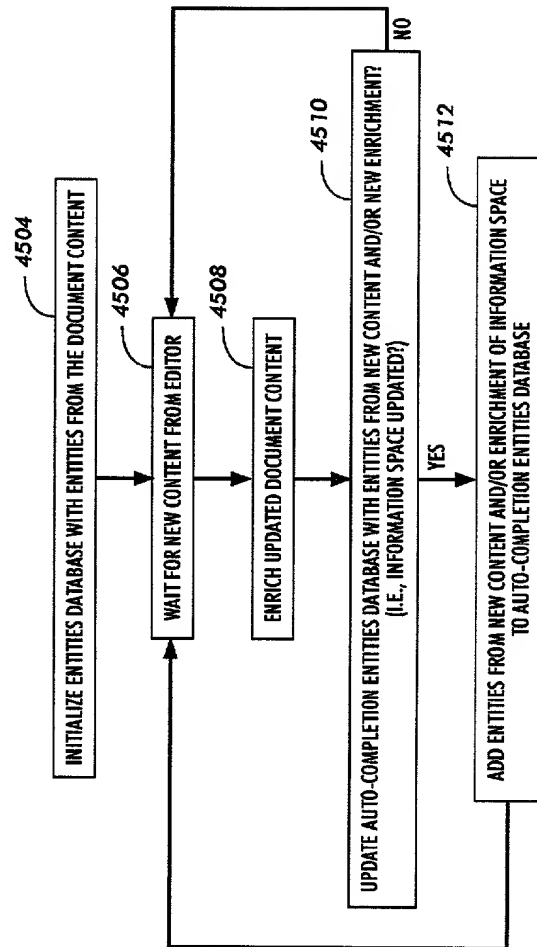


FIG. 45

46/67

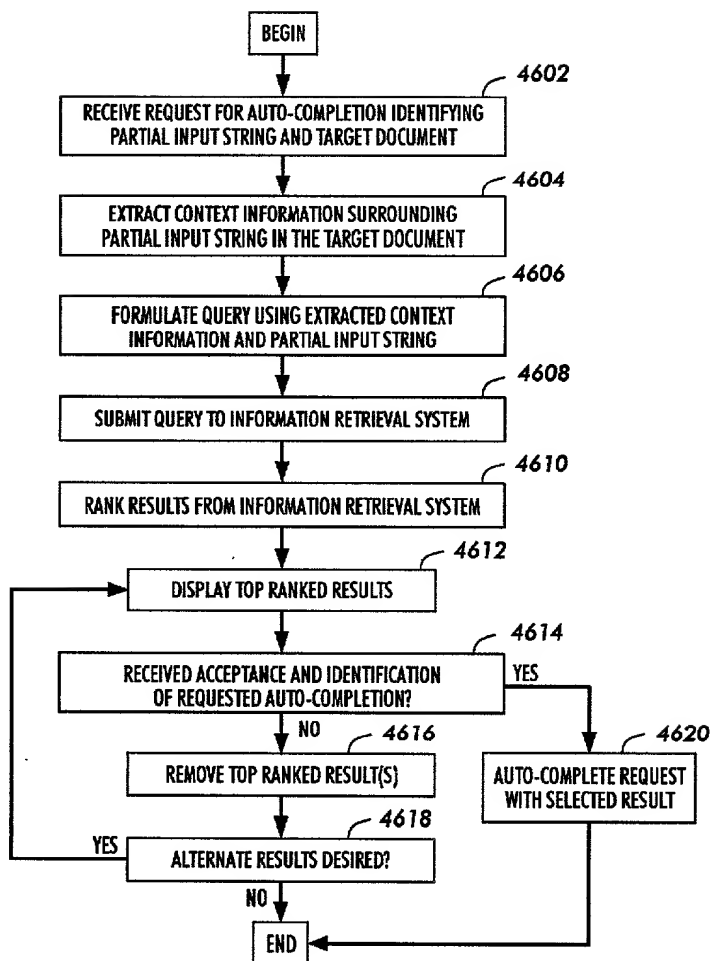
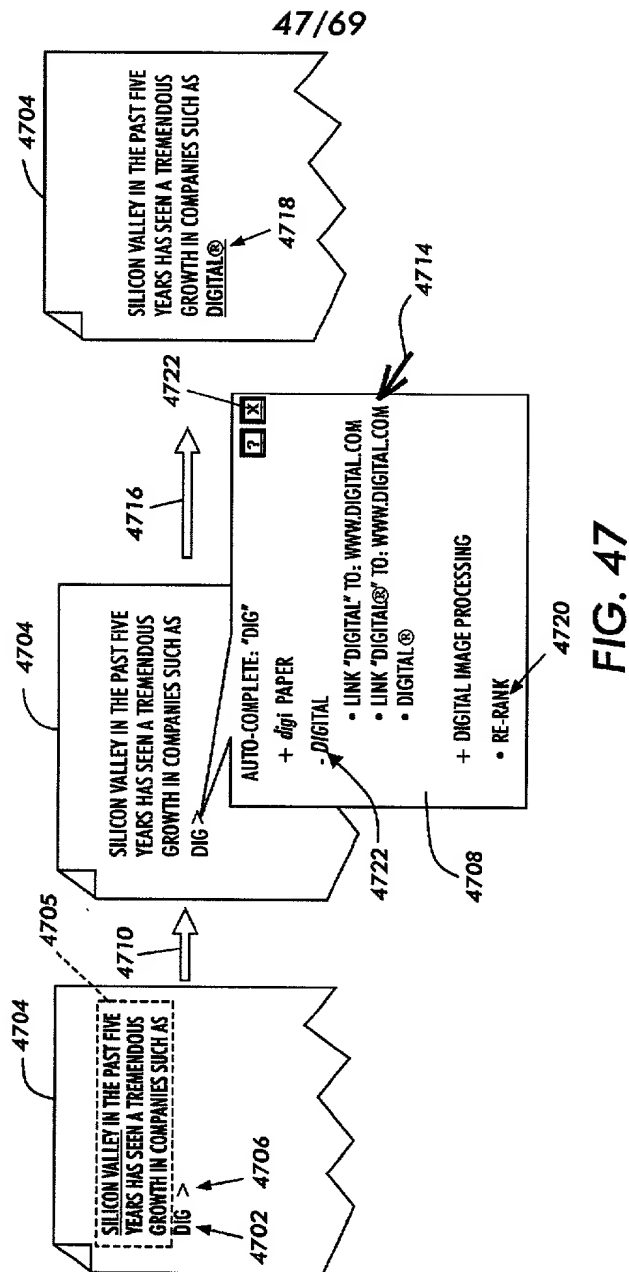


FIG. 46



48/69

(A) KEY	(B) ENTITY	(C) MARKUP	(D) FORMAT	(E) ORIGIN	(F) ORIGIN POSITION	(G) PART OF SPEECH	(H) CONTEXT
Sili	Silicon Valley	www.siliconvalley.com	Capitalize	TARGET DOCUMENT CONTENT	1	Noun	Science
Dig	Digital	www.digital.com	®	www.digital.com	434	Adjective	Science
Dig	DigiPaper	NONE	Bold	www.parc.xerox.com	23	Noun	Science
Dig	digress	NONE	NONE	www.nytimes.com	343	Verb	Politics
...

FIG. 48

49/69

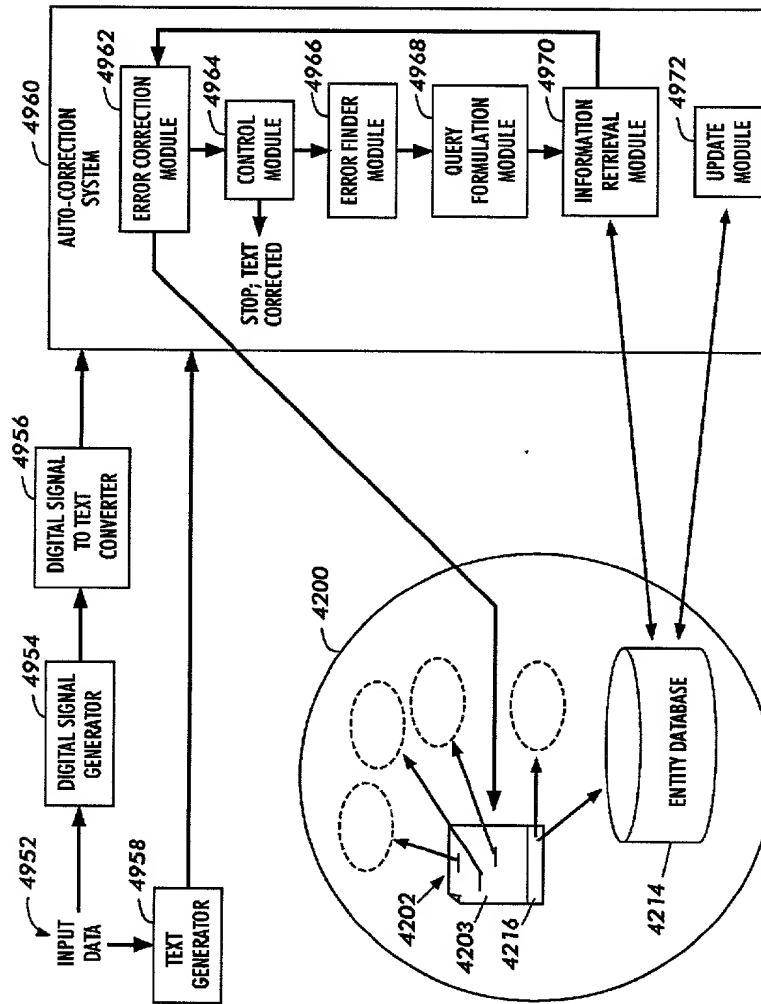


FIG. 49

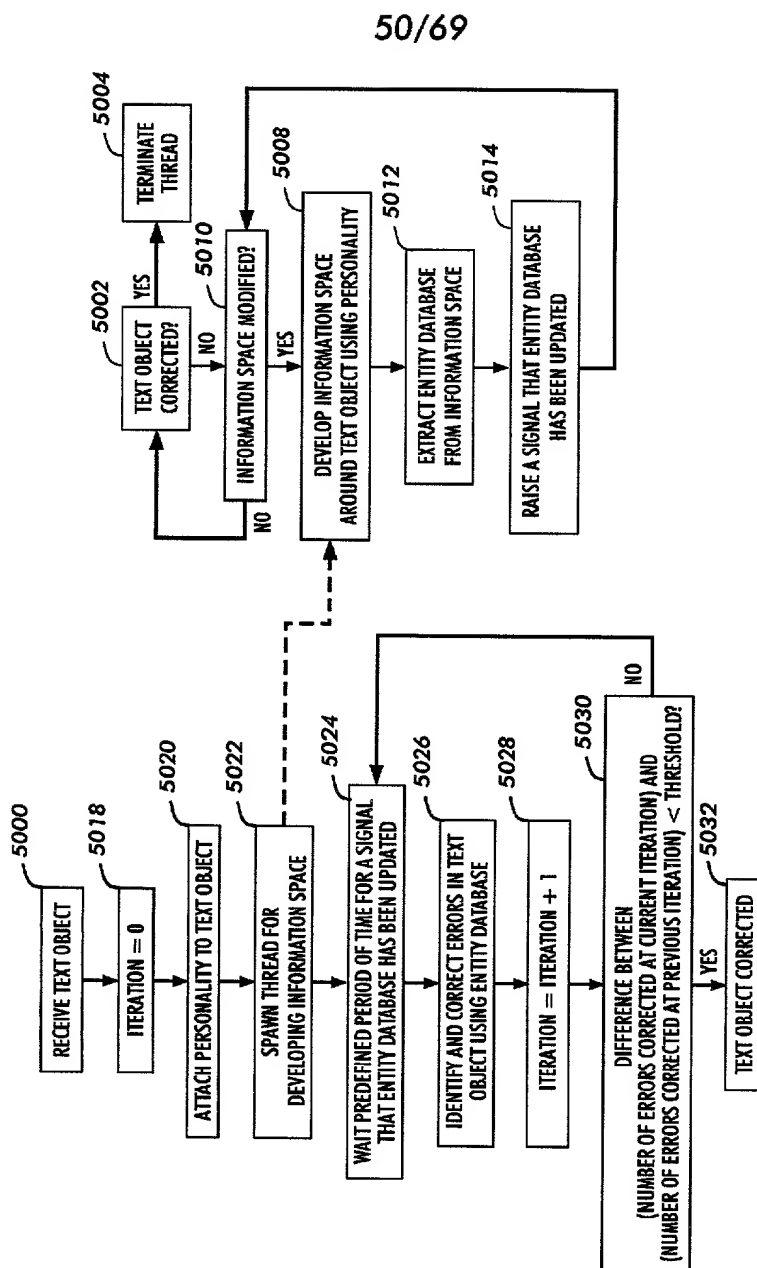


FIG. 50

51/69

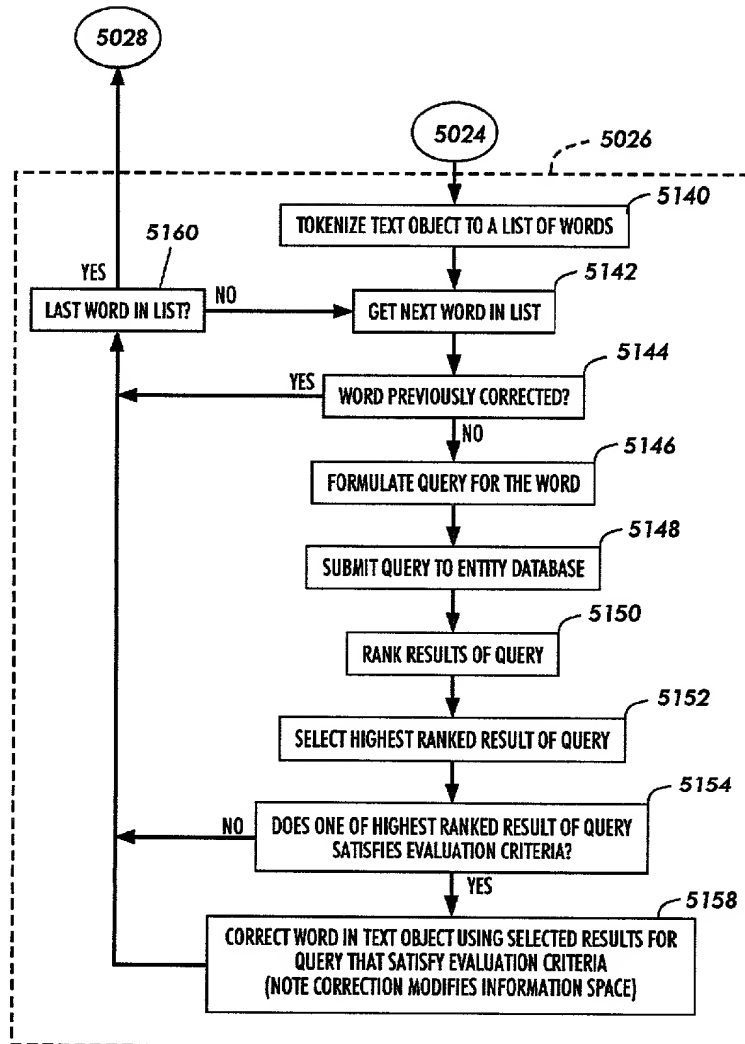


FIG. 51

52/69

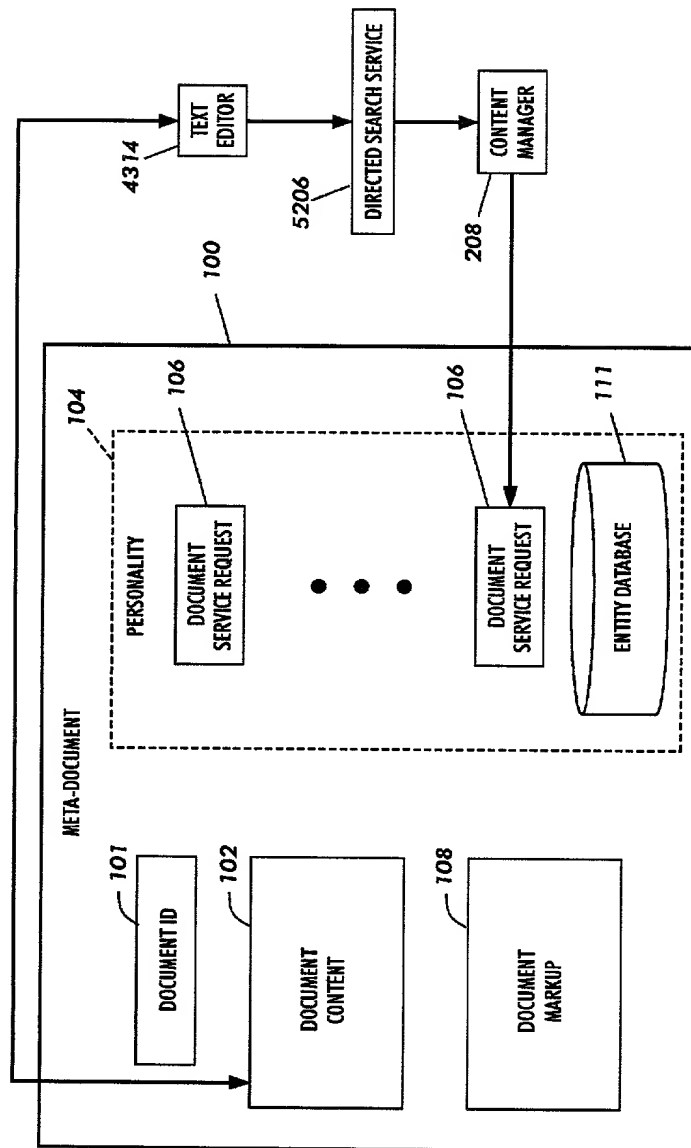


FIG. 52

53/69

5300 DIRECTED SEARCH

5302 Get: **5304** For:

5306 Insert Result as: ☐ Link ☒ Content ☐ Metadata

5308 Update Frequency:

5310 Select Providers:

- ☒ Reuters
- ☒ Thomson
- ☒ Hoover

5312 Ensures:

Language of text is:

☒ Premium ☐ Non-Premium

☐ Low Quality ☒ High Quality

☐ Copyrighted ☒ Not Copyrighted

☐ Free ☒ Paid Max price \$:

Maximum number of words:

5318 Search Selected Providers By:

5320 ☒ Cost

5322 ☐ Speed

5314 Location Of Results:

☒ After Current Selection

☐ Before Current Selection

☐ As Comment To Current Selection

☐ As Footnote To Current Selection

5315 Current Selection Is:

5317 Xerox in the past five years has seen tremendous growth

5316

FIG. 53

54/69

5400

```
<service id='Products'
  refresh=30
  insert='Content'
  language='En'
  quality='Premium&High&Not-copyrighted&paid'
  maxprice='50'
  maxwords='2000'
  search='cost'>
  <entity>Xerox
    <offsets>13 18 456 461 710 715</offsets>
    <provider name='Thomson' id=18></provider>
    <provider name='Reuters' id=19></provider>
  </entity>
</service>
```

FIG. 54

55/69

5500

IMPORT/EXPORT [?] [X]

EXPORT NAME: 5502

EXPORT:

- ☒ ORIGINAL DOCUMENT CONTENT
- ☒ DOCUMENT MARKUP (I.E., ENRICHMENT)
- ☒ PERSONALITIES

5504

5506

5508

☐ IMPORT NAME: 5512

IMPORT:

- ☒ ALL CONTENT
- ☐ SELECTED CONTENT ONLY
 - ☐ ORIGINAL DOCUMENT CONTENT
 - ☐ DOCUMENT MARKUP (I.E., ENRICHMENT)
 - ☐ PERSONALITIES

5510

- ☒ BIND UN-MATCHED ELEMENTS
- ☒ ENHANCE BOUND ELEMENTS WITH EXISTING SERVICE PROVIDERS

5514

FIG. 55

56/69

```

<!-- SERVER ID>
<docsouls:serverid> DS v1.0 server name </docsouls:serverid>
<!-- PROPERTIES OF DOCUMENT>
  <docsouls:docid> xxx </docsouls:docid>
  <docsouls:name> xxx </docsouls:name>
  <docsouls:owner> xxx </docsouls:owner>
  ...
<!-- ORIGINAL DOCUMENTS DOCUMENT CONTENT>
<docsouls:originaldoc format="HTML">toto.html</docsouls:originaldoc>
<docsouls:browsingdoc format="HTML">browse.html</docsouls:browsingdoc>

<!-- LIST OF PERSONALITIES AND DOCUMENT SERVICE REQUESTS>
<docsouls:endowmentdescription>
  <!-- LIST OF PERSONALITIES>
    <docsouls:personality id = " " name=" " etc.>
    <docsouls:service id=" " name=" " etc.>
    <docsouls:provider name=" " etc.>
    <docsouls:provider name=" " etc.>
    </docsouls:service>
  </docsouls:personality>
  ...
  <!-- LIST OF SERVICE REQUESTS>
    <docsouls:service id=" " name=" " etc.>
    <docsouls:provider name=" " etc.>
    <docsouls:provider name=" " etc.>
    </docsouls:service>
  ...
</docsouls:endowmentdescription>

```

FIG. 56

57/69

```

<!-- DESCRIPTIONS OF PERSONALITIES>
<docsouls:personalitydescription>
  <docsouls:id> xx </docsouls:id>
  <!-- all the fields AND service categories>
  <docsouls:servicecategory id="" name = "">
</docsouls:personalitydescription>

<!-- NEXT PERSONALITY DESCRIPTION>
...

<!-- DESCRIPTION OF DOCUMENT SERVICE REQUESTS>
<docsouls:servicedescription>
  <docsouls:id> xx </docsouls:id>
  <!-- list all fields (e.g., "refresh period", category, etc.)>
  <!-- list all providers>
  <docsouls:provider id="" name = "">
  <!-- list all strategies>
  <!-- list status of services (original, frozen, mapped, etc.)>
  <docsouls:servicestatus status="mapped" originalid="" originalserver="">
  <!-- list Serverid if the service was not available on current server>
  <docsouls:serverid>id</docsouls:serverid>
</docsouls:servicedescription>

<!-- NEXT SERVICE REQUESTS DESCRIPTION>
...

```

FIG. 57

58/69

```

<!-- DESCRIPTION OF PROVIDERS >
<docsouls:providerdescription>
  <docsouls:id> xx </docsouls:id>
  <!-- all the fields e.g. key, URL, etc.>
</docsouls: providerdescription >

<!-- DESCRIPTION OF DICTIONARIES >
<docsouls:dictionarydescription>
  <docsouls:id> xx </docsouls:id>
  <!-- all the fields, e.g. category>
</docsouls:dictionarydescription>

<!-- DESCRIPTION OF STRATEGY >
<docsouls:strategydescription>
  <docsouls:id> xx </docsouls:id>
  <!--all the fields>
  <!--all the dictionaries>
</docsouls:strategydescription>

<!-- RESULTS OF GLOBAL SERVICES >
<docsouls:globalservice id="Summary">
  <docsouls:link> xxx </docsouls:link>
  <docsouls:content> xxx </docsouls:content>
</docsouls:globalservice>

<!-- OTHER RESULTS (E.G. KEYWORDS, CATEGORIES, ETC.) >

```

FIG. 58A

59/69

```

<!-- (OPTIONAL) DESCRIPTION OF ENRICHMENT PERFORMED AT OTHER DOCUMENT ENRICHMENT
SYSTEMS>
<docsouls:serverid>olderServerId</docsouls:serverid>
<docsouls:personalitydescription>...</docsouls:personalitydescription>
<docsouls:servicedescription>...</docsouls:servicedescription>

<!-- (OPTIONAL) LIST OF RESULTS OF THE LAST ENRICHMENT>
<!-- IDENTIFY DOCUMENT SECTION IN CASE OF MULTI-PART DOC>
<docsouls:docpart>main_frame.html</docsouls:docpart>
<!-- OFFSET WHEN ENRICHMENT APPLIES TO A SECTION>
<docsouls:offset> 0 200 </docsouls:offset>
<!-- STATUS OF SERVICE APPLIED ("ACTIVE" OR "FROZEN")>
<docsouls:service id="stockQuote" status="active">
  <!-- ENTITY>
  <docsouls:entity>Xerox
    <!-- NORMALIZED FORM>
    <docsouls:normform>Xerox</docsouls:normform>
    <!-- LIST OF OFFSETS WHERE ENTITY IS FOUND IN TEXT RELATIVE TO SECTION>
    <docsouls:offsets start=35 end=40 />
    <docsouls:offsets start=232 end=237 />
    <!-- DICTIONARIES>
    <docsouls:dict> xxx </docsouls:dict>
    <!-- STRATEGIES>
    <docsouls:strat> xxx </docsouls:strat>
    <!-- LINK INFORMATION>
    <docsouls:link> xxx </docsouls:link>
    <!-- CONTENT INFORMATION>
    <docsouls:content> xxx </docsouls:content>
  </docsouls:entity>
  <!-- NEXT ENTITY>
  ...
</docsouls:service>

```

FIG. 58B

60/69

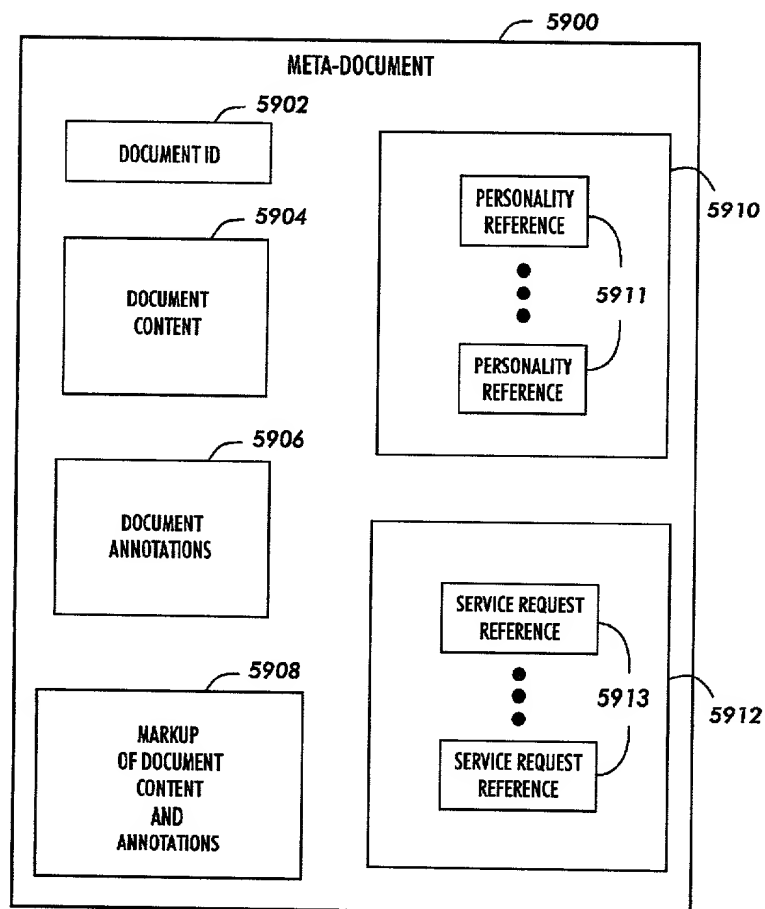


FIG. 59

61/69

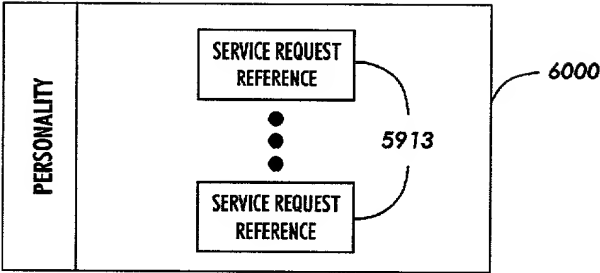


FIG. 60

62/69

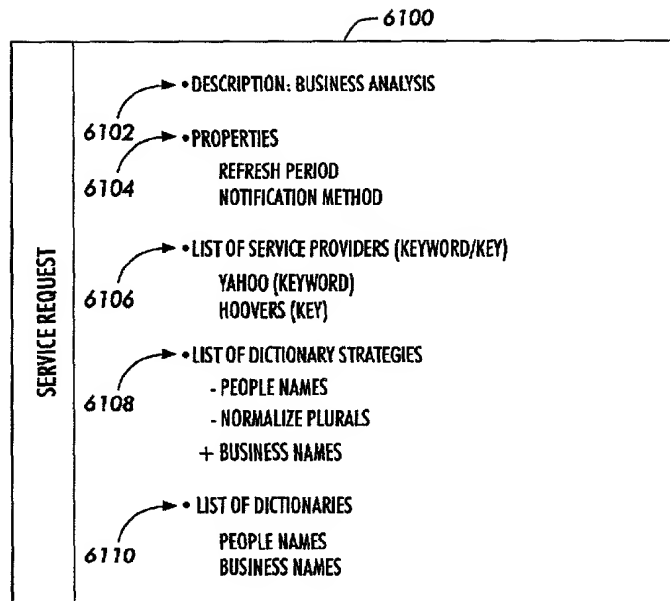
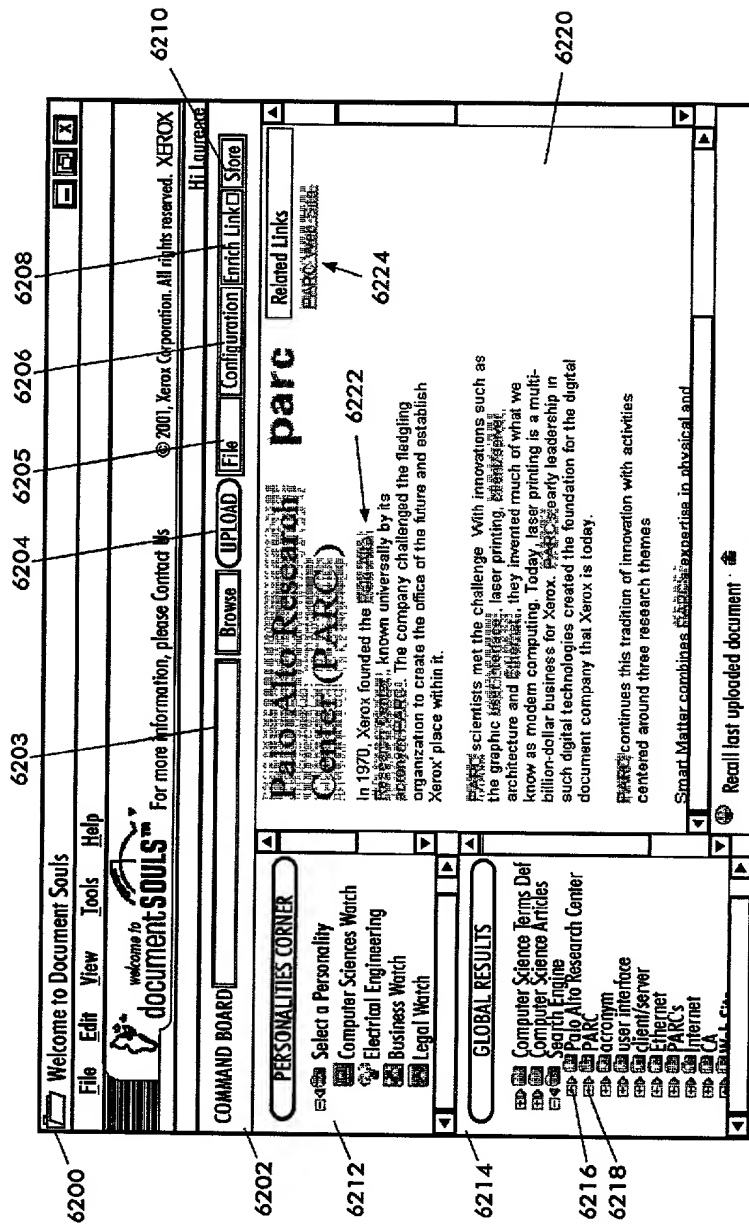


FIG. 61

63/69



64/69

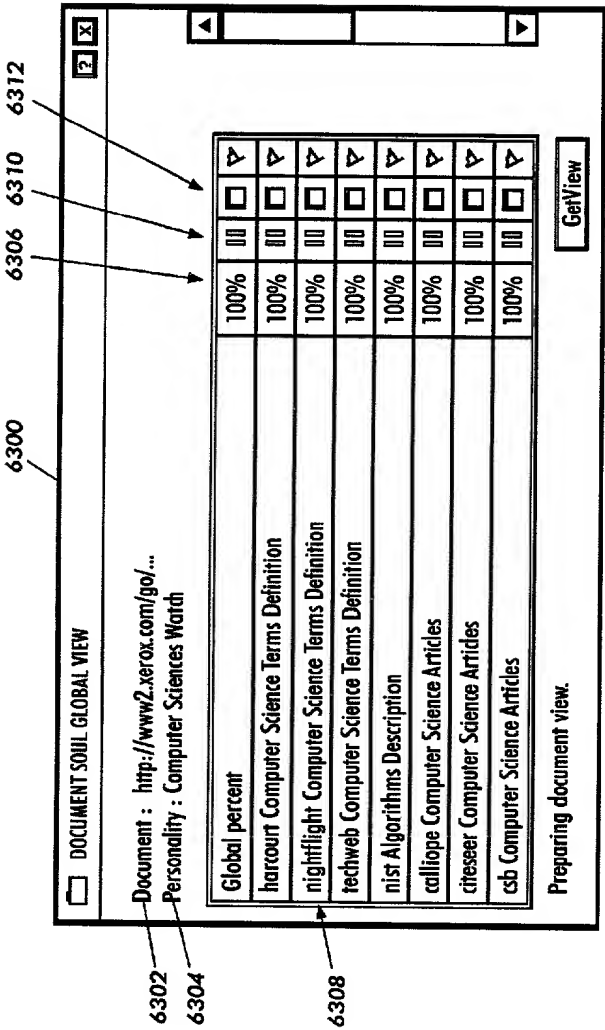


FIG. 63

65/69

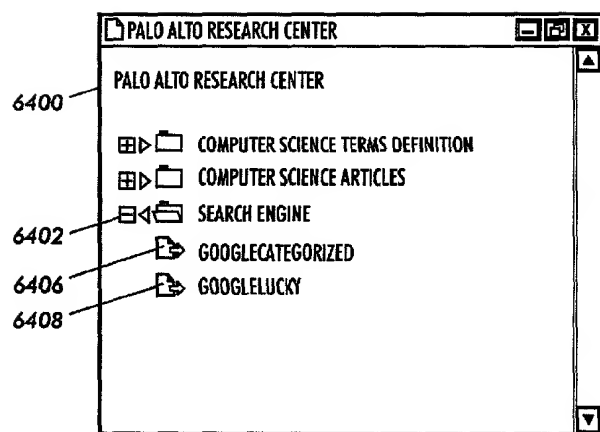


FIG. 64

66/69

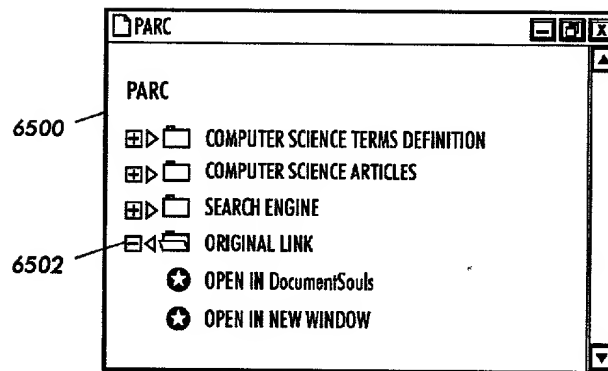


FIG. 65

67/69

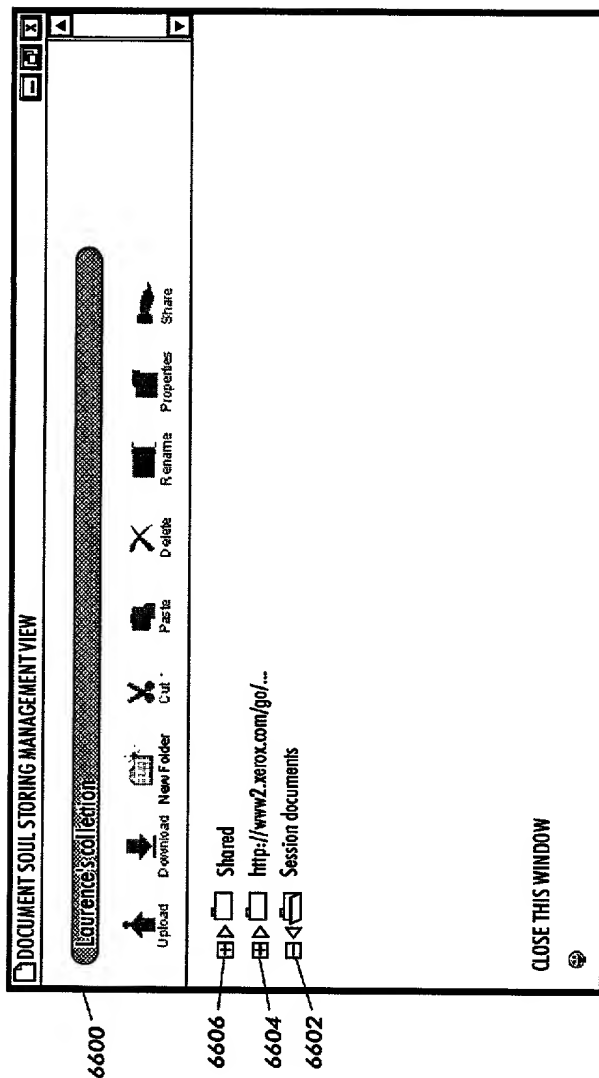


FIG. 66

68/69

DOCUMENT SOUL MARKING OPTIONS

Marking and Enrichment Options

Add an enrich button to your browser: Enrich with DocSoulz (drag & drop this link to your toolbar, Unix Macscape user: right click on the link and choose add bookmark for link then edit your bookmark and rename it)

Default personality:

- ☐ Choose default personality: Computer Science March
- ☐ Always ask for personality
-
- ☐ Automatically enrich any uploaded document with the default personality
- ☒ Enrich followed links

SUBMIT CANCEL

6700 6702 6704 6706 6708 6710

FIG. 67

69/69

[illegible]

FIG. 68